VOL. XXVI APRIL, 1935 NO. 4

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

CONTENTS

The Definition and Measurement of "g" (General Intelligence) . . . 241
GODFREY H. THOMSON

Multiple Factor Analysis and Psychological Concepts 263
S. N. F. CHANT

Personality Traits by Fiat. I. The Analysis of the Total Trait Scores and Keys of the Bernreuter Personality Inventory 273
IRVING LORGE

Chance and Personality Inventory Scores. 279
ROBERT G. BERNREUTER

Some Observations on Multiple-factor Analysis 284
OLIVE R. RUSSELL

Some Attitudes of Students in the R. O. T. C.. 291
HERBERT W. ROGERS

A Correlational Method Applicable to the Study of the Time Factor in Intelligence 307
M. E. ODOROFF

The Reliability of Increments. 312
FRANK K. SHUTTLEWORTH

Book Reviews 314

$6.00 per Year · Published Monthly September to May

WARWICK & YORK, INC.

BALTIMORE, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879; additional entry as Second Class Matter at York, Pa.

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

Established 1910

EDITORS

JACK W. DUNLAP
Fordham University

HAROLD E. JONES
University of California

PERCIVAL M. SYMONDS
Teachers College,
Columbia University

H. E. BUCHHOLZ, *Managing Editor*

THE price of the Journal is $6.00 a year in the United States; $6.40 in foreign countries. Part-year subscriptions are 90 cents for each number ordered. Back volumes are $7.00 each; back issues are $1.10 each except when more than five years old, and then $1.20 each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of the next succeeding number.

Unsolicited manuscripts should be accompanied with return postage. Manuscripts, books and other materials for review, and correspondence regarding editorial and business matters should be addressed to the Journal.

WARWICK AND YORK · *Publishers* · BALTIMORE, MD

An Experimental Study of Certain Factors Affecting Transfer of Training in Arithmetic

BY JAMES ROBERT OVERMAN

An investigation of the effect of instruction on three types of examples in two-place addition upon the pupils' ability to handle closely related types in addition and subtraction, and to determine whether the amount of transfer is a function of the method of teaching. The problem was to ascertain if the amount of such transfer can be increased by helping the pupils consciously to generalize the process and to formulate, from the type taught, a general method of procedure applicable to related types.

$3.00 *plus* 12¢ *postage*

WARWICK AND YORK, *Publishers*, BALTIMORE, MD.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXVI | April, 1935 | Number 4

THE DEFINITION AND MEASUREMENT OF "*g*" (GENERAL INTELLIGENCE)

GODFREY H. THOMSON

Moray House, University of Edinburgh

The main object of the present paper is to elucidate further the nature both of the indeterminate variate i in the measurement of g and of g itself. The paper is entirely theoretical, and it is assumed throughout that experimental errors and population-sampling errors are nil, the whole population of persons being accurately tested. An orthogonal matrix will be arrived at which will transform any set of values for g and the specifics, possessed by a given individual, into any number of different sets of values for g and the specifics which will give exactly the same scores in a *hierarchical* set of tests, so that we do not know with certainty which set of values the individual really possesses though we can write down a set which is the most probable, under certain assumptions; and by a proper addition of *singly-conforming* tests we can arrive at an exact measure of the g defined by this particular hierarchy, which g is a function of the first three tests chosen in the hierarchy.

I. ELEMENTARY CONSIDERATIONS

The fundamental reason why g cannot be determined exactly from a *hierarchical* set of tests (unless one of them is pure g) is that in such a set there is always one more uncorrelated component or "factor" than the number of tests: For each test has its own specific "factor," and in addition there is the general "factor." We can therefore assign any value we like to one of the components (say to g), and still have as many unknowns left as there are equations, so that we can solve for the remaining components. If for example in four tests a certain individual obtains the scores z_1, z_2, z_3, z_4, and if these tests are hierarchical, we can write

$$\left.\begin{aligned} z_1 &= l_1 g + m_1 s_1 \\ z_2 &= l_2 g + m_2 s_2 \\ z_3 &= l_3 g + m_3 s_3 \\ z_4 &= l_4 g + m_4 s_4 \end{aligned}\right\} \qquad (1)$$

where the z's are known for this individual, and where the l's and m's are calculated from the correlations. We have therefore four equations for the g, s_1, s_2, s_3, s_4 possessed by this individual, and we can assign him any value we like for g and still satisfy the equations by suitable values of s_1, s_2, s_3 and s_4.

But although we can thus assign any value we like to the g of one individual, we cannot properly proceed to do this for every individual of the population: For after we had done so we might find that the g's and s's we had assigned were correlated with one another, contrary to hypothesis. Hence the need of the more complicated considerations of the following sections.

Before we leave the simple considerations of the present section we can illustrate two other points. In the first place, the above equations make it clear that if we assign to an individual a higher g we must assign to him lower values in the specifics s, and *vice versa*. The indetermination of g is the same thing as the difficulty of knowing whether a man has scored high marks because he has a high g or because he has high specifics.[1] And in the second place it is clear that we can only do away with this lack of certainty either (*a*) by adding to our set of tests one which does *not* add a new component or factor[2] or (*b*) more generally, by adding x new tests which only depend upon $x - 1$ factors, in addition to those already existing.[3] We turn now to a more elaborate discussion of these various points.

II. NOTATION

Let z represent a set of k variates $\{z_1 z_2 z_3 \ldots z_i \ldots z_k\}$ (*e.g.* test scores) each variate being measured from its mean and in a unit such that for each test i, $\Sigma z_i^2 = 1$, the sum being over the population p of

---

[1] See Thomson: "The Meaning of i in the Estimate of g." *Brit. Journ. of Psychol.*, Vol. XXV (1), July, 1934, in Section IV.

[2] This can be one which depends on g alone and has no specific, or one which does not depend on g at all, but only on one of the already existing specifics, or one which depends on g and one of the already existing specifics in new proportions. Thomson, G. H.: "On Measuring g and s by tests which break the g-hierarchy." *Brit. Journ. Psychol.*, Vol. XXV (2), Oct., 1934, p. 204.

[3] Thomson, G. H.: *Nature*, January 12, 1935, page 71.

person's tested.[1] Measuring in such units from the mean will be called standardizing to unity. In *practical* standardizing, we make $\Sigma x_i^2 = p$: But in theoretical work it is more convenient to use a unit $\sqrt{p}$ times larger than the standard deviation, so that $\Sigma z_i^2 = 1$. There is no loss of generality in these changes, which are changes of unit only.

We can suppose that each z_i is the weighted sum of a number n of uncorrelated standardized components s. This is an assumption only if we attach some degree of reality to the components, in which case we are assuming that each z_i is a weighted *sum* (and not some other function) of the components, or "factors" as psychologists term them. Mathematically however it is not an assumption but merely a method of expression in terms of new variables.

We can then give a geometrical meaning to the z's, by regarding the uncorrelated components s as measured along a set of n rectangular axes. Each z_i is then represented by a line through the origin of this space, a space which has as many dimensions as there are components. Any person is represented by a point in this space, his scores in the k tests z are represented by the projections of his distance from the origin on to each of the z-lines in turn, and his outfit of "factors" is represented by the rectangular coordinates of his point.

The correlations between the variates are then represented geometrically by the cosines of the angles between the z-lines.

In matrix notation the relation between the z's and the components s is given by

$$z = Ms \tag{2}$$

which, written out at length for (say) three variates, and (say) five components, is

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ b_1 & b_2 & b_3 & b_4 & b_5 \\ c_1 & c_2 & c_3 & c_4 & c_5 \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} \tag{3}$$

where for reasons of later convenience we have preferred to number the components 0, 1, 2, . . . instead of 1, 2, 3. . . . The condition that each z and each s is in standard measure carries with it the require-

[1] Each individual of the population tested has of course his own score z_i, and a second subscript representing the individual may be understood: But we shall not at present need to print this second subscript.

ment that in each row of the matrix M the sum of the squares must be unity. The equation (3) means in ordinary notation the set of equations

$$\left.\begin{aligned} z_1 &= a_1 s_0 + a_2 s_1 + a_3 s_2 + a_4 s_3 + a_5 s_4 \\ z_2 &= b_1 s_0 + b_2 s_1 + b_3 s_2 + b_4 s_3 + b_5 s_4 \\ z_3 &= c_1 s_0 + c_2 s_1 + c_3 s_2 + c_4 s_3 + c_5 s_4 \end{aligned}\right\} \tag{4}$$

and the matrix equation (2) similarly means a set of k such equations. A matrix will be represented by a Roman capital, and a single column of a matrix by a small Roman letter. M' means the matrix M transposed, *i.e.* with its rows and columns interchanged, and similarly z' means the quantities z written in a row instead of in a column. Matrices are multiplied like determinants but always row-by-column: A matrix product MN does not necessarily or usually equal NM.

The product $MM' = R$ is the square matrix of correlations, *e.g.* for three variates

$$MM' = R = \begin{bmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{bmatrix} \tag{5}$$

If the matrix R is such that the determinants of all its two by two minors, excepting those involving a member of the principal diagonal, are zero—that is, if its tetrad-differences are zero—the set of variates producing it is said to be hierarchical, the matrix R itself may be called hierarchical, and we shall write it H. Such a hierarchical matrix $R = H$ will arise when the matrix M of equation (2) is of the special two-factor form L where

$$L = \begin{bmatrix} l_1 & m_1 & . & . & . \\ l_2 & . & m_2 & . & . \\ . & . & . & . & . \\ l_k & . & . & . & m_k \end{bmatrix} \quad (l_i^2 + m_i^2 = 1) \tag{6}$$

for in this case direct multiplication shows that

$$R = LL' = \begin{bmatrix} 1 & l_1 l_2 & l_1 l_3 & . & l_1 l_k \\ l_1 l_2 & 1 & l_2 l_3 & . & l_2 l_k \\ l_1 l_3 & l_2 l_3 & 1 & . & l_3 l_k \\ . & . & . & . & . \\ {}_k l_1 & l_2 l_k & l_3 l_k & . & 1 \end{bmatrix} = H \tag{7}$$

in which it is obvious that the tetrad-differences are zero. In the matrix L there are $k + 1$ columns, and when we write $M = L$ in equation (2) we are postulating $n = k + 1$ uncorrelated components, or in geometrical language $k + 1$ rectangular axes. One of these components (s_0) occurs in every z_i equation, it is a "general factor" or g. The others occur once each, they are "specific factors." In geometrical imagery, each z-line is *in* one of the planes defined by the s_0 axis taken with each of the k others in turn, none of the z-lines is in the space between these planes.

The converse of the proposition embodied in equations (6) and (7)—of the proposition that if (6) then (7)—needs stating with care. If (7) is true, that is if R is hierarchical, then the variates z *can* be expressed by a matrix equation.

$$z = Ls \tag{8}$$

But there is no necessity or compulsion about this. For we can always write

$$s = At \tag{9}$$

where A is any orthogonal matrix of $k + 1$ rows, or any slab of $k + 1$ rows of a larger orthogonal matrix; and then

$$z = (LA)t \tag{10}$$

where the variates z are expressed in terms of new uncorrelated components t, but *not* necessarily in such a way that only two t's come into each z_i equation.

This orthogonal change corresponds geometrically to rotations of the rectangular axes of reference to new positions (also rectangular) with the same origin, leaving the z-lines and the angles between them unaltered. When the number of components t is larger than the number of components s, the axes have been rotated out of the space formerly employed, and one or more new dimensions have to be added: As though I turned x and y axes out of the plane of this paper into a third dimension, and had to add a z axis in order to describe events occurring in this plane. The test scores still give the same correlations as before, for

$$\text{new } H = (LA)(LA)' = LAA'L' = LIL' = LL' = \text{old } H \tag{11}$$

where I is the unit matrix $[a_{ij}]$, $a = 1$ for $i = j$, 0 for $i \neq j$.

The number of components t can also be one *less* than $k + 1$ the number of components s, in which case the rotation has stopped at a

point where one of the axes is at right angles to the k-space defined by the k tests, and need not therefore be used.

Lastly, let us write for each test of a hierarchical set

$$q_i = \frac{l_i}{m_i} \tag{12}$$

and let q without any subscript mean the column vector

$$\{q_0 \; q_1 \; q_2 \; \ldots \; .\}$$

where $q_0 = -1$, and where the number of members in q is not at present stated. Then

$$q'q = 1 + \frac{{l_1}^2}{{m_1}^2} + \frac{{l_2}^2}{{m_2}^2} + \cdots \tag{13}$$

while qq' is a square matrix

$$qq' = \begin{bmatrix} 1 & -q_1 & -q_2 & -q_3 & . \\ -q_1 & {q_1}^2 & q_1q_2 & q_1q_3 & . \\ -q_2 & q_1q_2 & {q_2}^2 & q_2q_3 & . \\ -q_3 & q_1q_3 & q_2q_3 & {q_3}^2 & . \\ . & . & . & . & . \end{bmatrix} \tag{14}$$

III. THE TRANSFORMATION WHICH ALTERS THE "FACTORS" WHILE LEAVING ALL THE TESTS "TWO-FACTOR" IN NATURE AND THE CORRELATIONS UNCHANGED

Consider the matrix

$$B = I - \frac{2qq'}{q'q} \tag{15}$$

where I is the unit matrix. This matrix B is orthogonal, as can be readily checked, and is of unspecified size. Written out at length it is

$$B = \begin{bmatrix} 1-\lambda & \lambda q_1 & \lambda q_2 & \lambda q_3 & \lambda q_4 & . \\ \lambda q_1 & 1-\lambda {q_1}^2 & -\lambda q_1q_2 & -\lambda q_1q_3 & -\lambda q_1q_4 & . \\ \lambda q_2 & -\lambda q_1q_2 & 1-\lambda {q_2}^2 & -\lambda q_2q_3 & -\lambda q_2q_4 & . \\ \lambda q_3 & -\lambda q_1q_3 & -\lambda q_2q_3 & 1-\lambda {q_3}^2 & -\lambda q_3q_4 & . \\ \lambda q_4 & -\lambda q_1q_4 & -\lambda q_2q_4 & -\lambda q_3q_4 & 1-\lambda {q_4}^2 & . \\ . & . & . & . & . & . \end{bmatrix} \tag{15a}$$

where λ is an abbreviation for $2/q'q$.

This matrix B has the property that

$$L\bar{B} = L \tag{16}$$

(except that the new L may have some columns of zeros added on its right). Here $\bar{B}$ means the first $k + 1$ rows of B (since L has $k + 1$ columns), and the total number of columns and rows in B is at our disposal except of course that it must not be less than $k + 1$.

If, as an example, we write equation (16) out at length for only three tests (for economy of space) and for say a five column B we have

$$\begin{bmatrix} l_1 & m_1 & . & . \\ l_2 & . & m_2 & . \\ l_3 & . & . & m_3 \end{bmatrix} \begin{bmatrix} 1-\lambda & \lambda q_1 & \lambda q_2 & \lambda q_3 & \lambda q_4 \\ \lambda q_1 & 1-\lambda q_1^2 & -\lambda q_1 q_2 & -\lambda q_1 q_3 & -\lambda q_1 q_4 \\ \lambda q_2 & -\lambda q_1 q_2 & 1-\lambda q_2^2 & -\lambda q_2 q_3 & -\lambda q_2 q_4 \\ \lambda q_3 & -\lambda q_1 q_3 & -\lambda q_2 q_3 & 1-\lambda q_3^2 & -\lambda q_3 q_4 \end{bmatrix}$$

$$= \begin{bmatrix} l_1 & m_1 & . & . & . \\ l_2 & . & m_2 & . & . \\ l_3 & . & . & m_3 & . \end{bmatrix} \tag{17}$$

as actual multiplication confirms.

It will be noticed that it is quite immaterial what value we here give to q_4: But q_1, q_2, and q_3 are the l/m ratios from L.

In consequence of this property of the orthogonal matrix B we can, in the hierarchial set of tests $z = Ls$, write

$$s = \bar{B}t \tag{18}$$

and obtain

$$z = L\bar{B}t = Lt \tag{19}$$

where $t = \{t_0\ t_1\ t_2\ \ldots\}$, so that the tests z are now referred to a new set of components or factors, different from the former, of which t_0 is a new and different "general factor," and $t_1\ t_2\ t_3\ \ldots$ are new specific "factors." There may be any number of these latter, but if there are k tests z, then only $t_1\ t_2\ \ldots\ t_k$ will have weights different from zero, while t_{k+1} and all subsequent factors will have weight zero, *i.e.*, will not occur in the written out equations for the z's.

As a first numerical example take four tests which give the correlations shown below, which are hierarchical:

	1	2	3	4
1	.	.7483	.6928	.5657
2	.7483	.	.6481	.5292
3	.6928	.6481	.	.4899
4	.5657	.5292	.4899	.

(20)

The correlations of these four tests with g, calculated from these data, are

$$\begin{aligned} l_1 &= .8944 \text{ whence } m_1 = .4472,\ q_1 = 2.0000 \\ l_2 &= .8367 \qquad\quad\ \ m_2 = .5477,\ q_2 = 1.5275 \\ l_3 &= .7746 \qquad\quad\ \ m_3 = .6325,\ q_3 = 1.2247 \\ l_4 &= .6325 \qquad\quad\ \ m_4 = .7746,\ q_4 = 0.8165 \end{aligned}$$

If we write s_0 for g the equations for the four test-scores are

$$\left.\begin{aligned} z_1 &= .8944s_0 + .4472s_1 \\ z_2 &= .8367s_0 + .5477s_2 \\ z_3 &= .7746s_0 + .6325s_3 \\ z_4 &= .6325s_0 + .7745s_4 \end{aligned}\right\} \tag{21}$$

Here the "general factor" g or s_0 is measured along an axis which has the proper four angle $\cos^{-1} l$ with the four z-lines. But there are many other axes which have these angles with the four z-lines, other sets of rectangular axes in exact conformity with the correlations and which give "two-factor" equations. There is even one other set of axes fulfilling these conditions *in the same space as the s set.* For take the matrix B to be restricted to five rows and columns. In that case

$$\left.\begin{aligned} q &= \{\, q_0 \quad q_1 \quad\ q_2 \quad\ q_3 \quad\ q_4 \,\} \\ &= \{-1\ 2.0000\ 1.5275\ 1.2247\ 0.8165\}, \\ q'q &= 9.5,\ \lambda = .2105 \end{aligned}\right\} \tag{21a}$$

and the matrix B is (in two decimal places only to save space, reduced from six places)

$$B = \begin{bmatrix} .79 & .42 & .32 & .26 & .17 \\ .42 & .16 & -.64 & -.52 & -.34 \\ .32 & -.64 & .51 & -.39 & -.26 \\ .26 & -.52 & -.39 & .68 & -.21 \\ .17 & -.34 & -.26 & -.21 & .86 \end{bmatrix} \tag{22}$$

That is to say, the substitutions are

$$\left.\begin{aligned} s_0 &= .79t_0 + .42t_1 + .32t_2 + .26t_3 + .17t_4 \\ s_1 &= .42t_0 + .16t_1 - .64t_2 - .52t_3 - .34t_4 \\ s_2 &= .32t_0 - .64t_1 + .51t_2 - .39t_3 - .26t_4 \\ s_3 &= .26t_0 - .52t_1 - .39t_2 + .68t_3 - .21t_4 \\ s_4 &= .17t_0 - .34t_1 - .26t_2 - .21t_3 + .86t_4 \end{aligned}\right\} \tag{23}$$

Clearly s_0 is not the same thing as t_0, nor s_1 as t_1, etc. Yet when these substitutions (23) are made in the equations (21) they produce (*exactly* when all the decimals are kept) the equations.

$$\left.\begin{aligned} z_1 &= .8944t_0 + .4472t_1 \\ z_2 &= .8367t_0 + .5477t_2 \\ z_3 &= .7746t_0 + .6325t_3 \\ z_4 &= .6325t_0 + .7746t_4 \end{aligned}\right\} \tag{24}$$

that is to say, equations in t's with exactly the same coefficients as those of the s's in (21). These equations (24) produce the correlations (20) equally as well, obviously, as do equations (21). *We cannot tell whether g is s_0 or t_0.* Thus, for example, if we suppose that general intelligence and a large number of specific abilities are real entities and that a certain man possesses all of these to the extent of +1 standard deviation each, then he would score in the four tests, by equations (24), if we suppose the t's to be the proper axes, the scores

$$\left.\begin{aligned} z_1 &= 1.3416 \\ z_2 &= 1.3844 \\ z_3 &= 1.4071 \\ z_4 &= 1.4071 \end{aligned}\right\} \text{ standard deviations.}$$

But by equations (23) we see that we can alternatively ascribe to him

$$\begin{aligned} s_0 &= 1.96 \\ s_1 &= -.92 \\ s_2 &= -.46 \\ s_3 &= -.18 \\ s_4 &= .22 \end{aligned}$$

and these values inserted in equations (21) give the same four z-scores as above, so that we cannot tell whether his general intelligence is one standard deviation above the average ($t_0 = 1$) or nearly two ($s_0 = 1.96$). And we could make the change represented by equations (23) in every man in the population, and still have our "factors" uncorrelated with one another.

The "estimated" g of this man from this small battery of four tests would be, by the Spearman rules or by a regression equation,

$$\text{estimated } g = \frac{1}{1 + \sum \frac{l^2}{m^2}} \sum \frac{zl}{m^2} = 1.48$$

i.e., a value exactly half-way between the g we are supposing that he actually possesses (+1) and the g which his indistinguishable double possesses (1.96).

So much even if we remain in the same space. But if we begin to turn our axes into new dimensions, we can find not only two but innumerable sets which fulfill all the two-factor requirements and give the same correlations. Algebraically we do this by taking B of larger dimensions. Let us do this in the same numerical example, and let us add only one more dimension. That is to say, instead of the B of equations (22) made from the q of (21*a*) which was $\{q_0\ q_1\ q_2\ q_3\ q_4\}$ let us add a q_5. The former members of q were calculated from the data, but q_5 is entirely at our disposal. Let us for example take it so that $q_5^2 = 0.5$, so that $q'q = 10$ and $\lambda = 0.2$ exactly. Then the matrix B is (to two decimal places reduced from six)

$$B = \begin{bmatrix} .80 & .40 & .31 & .24 & .16 & .14 \\ .40 & .20 & -.61 & -.49 & -.33 & -.28 \\ .31 & -.61 & .53 & -.37 & -.25 & -.22 \\ .24 & -.49 & -.37 & .70 & -.20 & -.17 \\ .16 & -.33 & -.25 & -.20 & .87 & -.12 \\ .14 & -.28 & -.22 & -.17 & -.12 & .90 \end{bmatrix} \tag{25}$$

and the first four rows give us

$$\left.\begin{aligned} s_0 &= .80t_0 + .40t_1 + .31t_2 + .24t_3 + .16t_4 + .14t_5 \\ s_1 &= .40t_0 + .20t_1 - .61t_2 - .49t_3 - .33t_4 - .28t_5 \\ s_2 &= .31t_0 - .61t_1 + .53t_2 - .37t_3 - .25t_4 - .22t_5 \\ s_3 &= .24t_0 - .49t_1 - .37t_2 + .70t_3 - .20t_4 - .17t_5 \\ s_4 &= .16t_0 - .33t_1 - .25t_2 - .20t_3 + .87t_4 - .12t_5 \end{aligned}\right\} \tag{26}$$

If we substitute from equations (26) in equations (21) we again arrive at equations (24), *the quantity t_5 disappearing.* So we have here a new set of axes $t_0\ t_1\ t_2\ t_3\ t_4\ t_5$, of which, however, t_5 is at right angles to the space of the z-lines, though it is needed to describe the old axes in terms of the new. The space of the z-lines, a 4-space, remains of course the same, but the 5-space of the axes which contains it can be changed over and over again, just as, by analogy, many different planes can contain one line.

Our imaginary man who is one standard deviation above the average in general intelligence and in all specific abilities may have $t_0 = t_1 = t_2 = t_3 = t_4 = t_5 = 1$. But he may equally well have, from equations (26), values $s_0 = 2.05$, $s_1 = -1.11$, $s_2 = -0.61$, $s_3 = -0.29$,

$s_4 = 0.13$, *i.e.*, be indistinguishable from a man with $g = 2.05$, or if we take his t_5 (of which *ex hypothesi* we know nothing at all from these tests) to be, say, two standard deviations *below* the average, he would have $s_0 = 1.63$, $s_1 = -.27$, $s_2 = .05$, $s_3 = .22$, $s_4 = .49$. And so on. We have at our disposal the size of the matrix B, and the values of all the q's after those we know; we can turn the axes which represent g and the specifics about in a higher dimensioned space almost as we please, and indeed in any one space we pause in, there are two positions fulfilling all the conditions. We can almost visualize this last fact by reminding ourselves that it is the angle of the g axis with a z-line which has to remain constant: That this will be so if the g axis moves round the z-axis as the generating line of a cone: And that all these cones, as many as there are z-lines, will have one other common generating line in addition to the g axis we began with. A model which I have made in three dimensions lies before me as I write.

The rather extreme variations in the value of g which occur in the above small numerical example must not of course be made too much of, for they are large because there are only four tests in the battery and because the highest correlation of the four with g is .8944. With more tests, or a higher l among them, the indeterminate element in g becomes less. In such a case, $q'q$ becomes very large, and therefore $B = I - 2qq'/q'q$ becomes more and more closely $B = I$, the unit matrix, which leaves $s = It = t$. Nevertheless, even with as large a number of tests as seems reasonably practicable, some remarkable numerical results are produced. Suppose, for example, that we have an unlimited supply of tests each of which correlates with every other to the amount $r = .79839$, and each of which therefore correlates with g to an amount $l = .89353$. Suppose further that an individual actually possesses g and each and every specific to the amount of one standard deviation above the average. Then we get the following very remarkable series of estimates of his g by using ever more and more of our tests:

No. of tests	1	5	10	20	100	1000
Estimated g	(1.200)	1.430	1.466	1.484	1.499	1.502

where I have placed the value 1.200 in brackets since in a certain sense we can do nothing with less than four tests toward measuring g.

Here we actually have estimated g moving further and further away from its real value $+1$ as the number of tests is increased! The reason, of course, is that we are postulating a man who has $+1$ ability in each new specific that we add. Such a man is indistinguishable on these tests from one who has only average specifics and $g = 1.503$, and it is this last value which our estimate is climbing up towards.

IV. THE EQUATION OF THE ARRAY OF PERSONS WHO ALL SCORE ALIKE IN A HIERARCHICAL SET OF TESTS

If we think of a test as represented by a z-line through the origin O of a set of rectangular axes along which g, s_1, s_2, s_3 . . . s_k are measured, then, as we said above, each person j is represented by a point P in this $(k + 1)$-space, and his score z_{ij} in test i is represented by the projection of OP on the line z_i. All the persons whose points are in the k-space at right angles to z_i at that point z_{ij} will have the same score z_{ij} as he has. For example, in two dimensions (g and s_1) this k space is another line at right angles to z_1 through the given point: In three dimensions (g, s_1, and s_2) it is a plane. If there are two tests, there are two such planes, and along their common line the points represent persons who all have exactly the same two scores, z_1 and z_2. Similarly if there are k hierarchical tests, in the $(k + 1)$-space of g and the specifics there are k subspaces, each of k dimensions, and each at right angles to one of the lines at a given point. These k subspaces intersect along a line, every point on which represents a person with a different outfit of g and specifics, but with the same set of test scores, and all indistinguishable by these tests.

The equation of the space at right angles to the line z_i at the distance z_i from the origin is

$$z_i = l_i g + m_i s_i \tag{27}$$

where z_i, l_i, and m_i are constants and g and s_i are current coordinates. The set of equations which define the line along which the array of indistinguishable persons lies is the matrix equation

$$z = Ls \tag{28}$$

the same equation as (8). There, equation (8) was regarded as a means of finding test-scores given a man's g and specifics. Here equation (28) shows the relationship that must exist between a man's g and specifics if his test-scores are known. Given s and L we can find the point z. But given z and L we have a *line* $z = Ls$, not a

point. The reason is that the number of columns in L is one more than the number of rows. The indeterminateness of g will therefore persist until we make L square.

V. ABOLISHING THE INDETERMINATENESS OF g

We can make L square by adding a test which contains no new factor, but only g and one of the existing specifics.[1] The correlation of such a test with that already existing test whose specific it shares will break the hierarchy, but either of the two alone will fit into the hierarchy, and for each the l and m can be separately calculated from the hierarchy. If these two tests are called a and b, the condition that they should be expressible in terms of the same two factors g and s is

$$\Delta = \begin{vmatrix} 1 & l_a & l_b \\ l_a & 1 & r_{ab} \\ l_b & r_{ab} & 1 \end{vmatrix} = 0 \qquad (29)$$

and this at once shows that the determination of g has become exact, for the variance of g estimated from these two tests is Δ/Δ_{00} (where Δ_{00} is the minor which omits the row and column of l's) and therefore is zero.

Put in this way, this case at once suggests its own generalization. If we can find a number of new tests such that although no two of them can be added to the existing hierarchy without breaking it, yet each one can alone be so added, then for each of these tests we can find the correlation l with the g of the existing hierarchy. If then we find that

$$\Delta = \begin{vmatrix} 1 & l_a & l_b & l_c & . \\ l_a & 1 & r_{ab} & r_{ac} & . \\ l_b & r_{ab} & 1 & r_{bc} & . \\ l_c & r_{ac} & r_{bc} & 1 & . \\ . & . & . & . & . \end{vmatrix} = 0 \qquad (30)$$

then g can be estimated from these "singly conforming" tests with exactness[2] by the regression equation

---

[1] Thomson, G. H.: "On Measuring g and s by tests which break the g hierarchy." *Brit. Journ. of Psychol.*, Vol. XXV (2), Oct., 1934, p. 204.

[2] Thomson, G. H.: *Nature*, Jan. 12, 1935, p. 71. I there called such tests *non-conforming*, but *singly-conforming* seems a more accurate name.

$$\begin{vmatrix} g & z_a & z_b & z_c & . \\ l_a & 1 & r_{ab} & r_{ac} & . \\ l_b & r_{ab} & 1 & r_{bc} & . \\ l_c & r_{ac} & r_{bc} & 1 & . \\ . & . & . & . & . \end{vmatrix} = 0 \qquad (31)$$

In this case the "singly-conforming" tests, say p in number—"singly-conforming" because any one of them can enter the hierarchy, but any two of them break it—may be looked upon as adding only $p - 1$ new "specifics," or group factors. They are composed of p factors, namely g and $p - 1$ others. They break down the fatal excess of factors over tests of a pure hierarchy.

To prevent a misunderstanding which I have found not uncommon when teaching the above, I may add some further explanation of equation (30). Of course these or any p tests can always be expressed as linear functions of the same number p of components, or in geometrical language can be represented as radii vectores in a framework of p rectangular axes. If the l's be taken for the moment to represent the cosines of the angles which these radii make with any chosen one of the axes, then the determinant Δ will equal zero, from a well known proposition concerning the angles between $p + 1$ radii vectores in a p-space: For we have in effect added one of the axes to the number of radii vectores.

The point about our equation (30) is that the determinant equals zero *when the l's are calculated from the preexisting hierarchy.* It need not equal zero. It will only do so if there exist in the p-space of the singly-conforming tests a line which makes just these angles $\cos^{-1} l$ with the radii vectores representing the tests: That is if the g-line is in the same space as they are and not in a space of higher dimensions. But before we continue, it will be convenient at this point to introduce an abbreviated notation for matrices like those of equations (29) (30) and (31).

VI. MORE ON NOTATION

We have used R for any square matrix of correlations with unities in the principal diagonal, (see equation (5)), and H for such a matrix if it is hierarchical. Let us then write F for the matrix of correlations of the singly-conforming tests of equations (29) and (30). Equation 30 can then be abbreviated to

$$\begin{vmatrix} 1 & l' \\ l & F \end{vmatrix} = 0 \tag{32}$$

where l is a column and l' a row of the quantities l_a l_b. . . . Similarly equation (31) becomes

$$\begin{vmatrix} g & z' \\ l & F \end{vmatrix} = 0 \tag{33}$$

With such a notation, for a further example, the equation which gives g from a set of hierarchical tests is

$$\begin{vmatrix} g & z' \\ l & H \end{vmatrix} + i \begin{vmatrix} 1 & l' \\ l & H \end{vmatrix}^{\frac{1}{2}} = 0 \tag{34}$$

and i cannot disappear because $\begin{vmatrix} 1 & l' \\ l & H \end{vmatrix}$ is never zero (unless indeed one test is pure g). Fully written out, equation (33) would be

$$\begin{vmatrix} g & z' \\ l & F \end{vmatrix} + i \begin{vmatrix} 1 & l' \\ l & F \end{vmatrix}^{\frac{1}{2}} = 0 \tag{35}$$

but the determinant which is the coefficient of i is here, by equation (32), equal to zero.

It is illuminating to note that equations (32) and (33) can be enlarged to comprise all the tests, both the hierarchical set and the singly-conforming set. They would then become

$$\begin{vmatrix} 1 & l' \\ l & R \end{vmatrix} = 0 \tag{36}$$

and

$$\begin{vmatrix} g & z' \\ l & R \end{vmatrix} = 0 \tag{37}$$

where R is the matrix of correlation of *all* the tests, l a column of *all* the correlations with g, z' a row of a man's scores in *all* the tests. For if (32) is true, we can by Chio's method quickly prove that (36) is true also. In any actual experiment, it is presumably these larger sets of equations which will be used: And the experiment may be briefly described as finding first a set of sufficiently hierarchical tests, and then adding to them other tests which can only enter the hierarchy singly, but not in pairs, and which make the bordered determinant of equation (36) sufficiently near to zero. The necessity for

the additional tests being such as can enter the hierarchy singly arises from the necessity of calculating their correlations l with g.

VII. ON EXPRESSING ANY k TESTS BY k "FACTORS"

If we return now to our discussion of the possible misunderstanding referred to in the last two paragraphs of Section V, we may note that of course we could have made L square (see the last sentence of Section IV) by adding *any new test whatever* to the hierarchy, or indeed by rotating the axes of the existing hierarchy: But this would not have enabled us to measure with accuracy the g defined by the existing hierarchy.

If we add to the existing hierarchy any new test whatever, we can express this new test in terms of the g and specifics of the hierarchy without any new factors (though some of the coefficients may in some cases contain the root of minus one as factor) and L will then become a square matrix. The method of doing this is expressible very simply in our notation. Let the k hierarchical tests be $z = Ls$, and let the new test be $z_{k+1} = n's$ where $n' = [n_0\ n_1\ n_2\ \ldots\ n_k]$, so that the enlarged group of $k + 1$ tests is given by $z = Ns$, where the matrix N is square, and is made by adding the row of coefficients $n' = [n_0\ n_1\ n_2\ \ldots\ n_k]$ below L, *i.e.*

$$N = \begin{bmatrix} L \\ n' \end{bmatrix} \tag{38}$$

Let us further represent by $[c_0\ c_1\ c_2\ c_3\ \ldots\ c_k]$ the columns of N, so that for example c_0 is the whole column $\{l_1\ l_2\ l_3\ \ldots\ l_k n_0\}$, c_1 is the column $\{m_1\ 0\ 0\ \ldots\ n_1\}$ and so on. Then the quantities n_i can be found by solving the quadratic equations

$$\begin{vmatrix} 1 & c_i' \\ c_i & R \end{vmatrix} = 0 \tag{39}$$

where R is the matrix of correlations of the whole group of $k + 1$ tests. For n_0 this becomes, since the first k tests are hierarchical,

$$n_0^2\left(1 + \sum\frac{l^2}{m^2}\right) - 2n_0\sum\frac{lr}{m} + \sum\frac{r^2}{m^2} = 0 \tag{40}$$

where the summations are over the k hierarchical tests excluding the new test, and the r's are the correlations of the new test with them. After n_0 has been found, instead of solving the determinantal

quadratics for the other n's we can more easily find them from the equations

$$r_i = l_i n_0 + m_i n_i \tag{41}$$

where r_i is the correlation of the new test with test i of the preexisting hierarchy.

For example let the matrix R be

$$R = \begin{bmatrix} 1 & .669 & .566 & .438 & .560 \\ .669 & 1 & .592 & .458 & .554 \\ .566 & .592 & 1 & .387 & .566 \\ .438 & .458 & .387 & 1 & .721 \\ .560 & .554 & .566 & .721 & 1 \end{bmatrix} \tag{42}$$

where the first four tests are hierarchical and the fifth is a test added at hazard. From the first four tests, the correlation l of each with g can be found, and used in equation (40), from which, with equation (41), we then find the bottom row of the N matrix and obtain the two solutions

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} .80 & .60 & . & . & . \\ .84 & . & .55 & . & . \\ .71 & . & . & .71 & . \\ .55 & . & . & . & .84 \\ .40 & .40 & .40 & .40 & .60 \\ \text{or} & \text{or} & \text{or} & \text{or} & \text{or} \\ .87 & -.23 & -.32 & -.07 & .29 \end{bmatrix} \begin{bmatrix} s_0 \text{ or } g \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} \tag{43}$$

The matrix N with columns $[c_0\, c_1\, c_2\, \ldots\, c_k]$ having thus been found, g (it might then be thought) can be calculated (exactly, since N is square) from the equation

$$\begin{vmatrix} g & z' \\ c_0 & R \end{vmatrix} = 0 \tag{44}$$

with of course two solutions since there are two values for n_0. But while this is in a sense true, yet neither of these solutions for g is that which would be given by the hierarchical set of tests if they were increased in number till infinitely numerous. The finite hierarchy gives innumerable measures of g, and all that the addition to it of just *any* new test has done is arbitrarily to select two of these g's for preference.

In the case of our numerical example, equation (44) fully written out is

$$\begin{vmatrix} g & z_1 & z_2 & z_3 & z_4 & z_5 \\ .80 & 1 & .67 & .57 & .44 & .56 \\ .84 & .67 & 1 & .59 & .46 & .55 \\ .71 & .57 & .59 & 1 & .39 & .57 \\ .55 & .44 & .46 & .39 & 1 & .72 \\ .40 & .56 & .55 & .57 & .72 & 1 \end{vmatrix} = 0 \qquad (45)$$

or

.87

which will give a man's "g" quite exactly (except for the double solution) if his test scores $z_1z_2z_3z_4z_5$ are inserted: But it is a "g" which is a function of the test which has been added, and is not the g defined (though not measured) by the preexisting hierarchy.

But, it might next be erroneously thought, although making the matrix square by adding to the hierarchy just *any* new test results in measuring the "wrong" g, perhaps this will not be the case if the new test conforms to the hierarchy and is then expressed in terms of the preexisting factors, as can of course be done. This, it will be seen, is the same thing as asking whether we cannot measure g by making L square through rotating the s axes till one of them goes out of use, and that in such a way that only one row, say the last, of L is affected, the first k-1 rows being left unchanged.

VIII. THE TRANSFORMATION WHICH DESTROYS ONE OF THE SPECIFICS IN A HIERARCHICAL SET OF TESTS, LEAVING THE CORRELATIONS UNCHANGED AND ALL THE TESTS BUT ONE STILL "TWO-FACTOR"

The question is, whether given a set of hierarchical tests $z = Ls$ we can find a transformation $s = Dt$ such that in $z = LDt$ the matrix (LD) is identical with L in all its rows except the last, and has in its right hand column nothing but zeros so that the effective part of the matrix is square. That such a matrix D must exist is clear from the preceding section, by the methods of which we can find the effective square part of the matrix (LD). For example, if we have five tests and $z = Ls$ fully written out is

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} .800 & .600 & . & . & . & . \\ .837 & . & .548 & . & . & . \\ .707 & . & . & .707 & . & . \\ .548 & . & . & . & .837 & . \\ .400 & . & . & . & . & .9165 \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix} \qquad (46)$$

then Section VII enables us readily to find $z = Nt$. For we treat z_5 as *any* test added to a hierarchy of four tests and obtain

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} .800 & .600 & . & . & . & . \\ .837 & . & .548 & . & . & . \\ .707 & . & . & .707 & . & . \\ 548 & . & . & . & .837 & . \\ 702 & -.403 & -.462 & -.302 & -.198 & . \end{bmatrix} \begin{bmatrix} t_0 \\ t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix} \tag{47}$$

as well as another solution where the bottom row is $-.025$, $.566$, $.649$, $.425$, $.278$, 0, in both of which solutions t_5 has gone out of action; and we want to know if t_0 is the same thing as s_0. The very fact that there are two solutions shows that it is not: But it is of interest to see what the matrix D is which shows the connection between the s_0 of equation (46) and the t_0 of equation (47).

The method of finding D is to form N (as above in the numerical case) and then solve the equations for each member of D, which is rather easily done as certain quadratics which arise have for discriminants the equation (39) and are therefore perfect squares. There results

$$D = \begin{bmatrix} 1 \mp \mu & \pm \mu q_1 & . & \pm \mu q_{k-1} & \pm \nu \\ \pm \mu q_1 & 1 \mp \mu q_1^2 & . & \mp \mu q_1 q_{k-1} & \mp \nu q_1 \\ . & . & . & . & . \\ \pm \mu q_{k-1} & \mp \mu q_1 q_{k-1} & . & 1 \mp \mu q^2_{k-1} & \mp \nu q_{k-1} \\ \pm \nu & \mp \nu q_1 & . & \mp \nu q_{k-1} & \mp \nu q_k \end{bmatrix} \tag{48}$$

where μ is an abbreviation for $\dfrac{1}{Q(q_k \pm Q)}$

and ν means $\dfrac{1}{Q} = \dfrac{1}{+\sqrt{(1 + q_1^2 + q_2^2 + \cdots + q_k^2)}}$

In the case of our numerical example therefore we have

$$D_1 = \begin{bmatrix} .873 & .170 & .194 & .127 & .083 & .385 \\ .170 & .774 & -.259 & -.170 & -.111 & -.514 \\ .194 & -.259 & .703 & -.194 & -.127 & -.589 \\ .127 & -.170 & -.194 & .873 & -.083 & -.385 \\ .083 & -.111 & -.127 & -.083 & .945 & -.252 \\ .385 & -.514 & -.589 & -.385 & -.252 & -.168 \end{bmatrix} \tag{49}$$

$$D_2 = \begin{bmatrix} .821 & .238 & .273 & .179 & .117 & -.385 \\ .238 & .682 & -.364 & -.238 & -.156 & .514 \\ .273 & -.364 & .583 & -.273 & -.179 & .589 \\ .179 & -.238 & -.273 & .821 & -.117 & .385 \\ .117 & -.156 & -.179 & -.117 & .923 & .252 \\ -.385 & .514 & .589 & .385 & .252 & .168 \end{bmatrix} \quad (49a)$$

two matrices which, premultiplied by the matrix of equation (46) give respectively the two solutions of equation (47). They make it clear that s_0 and t_0 are quite different, so that the exact finding of t_0 does not give us the general factor which is defined by the hierarchy of equation (46) and could be measured by extending that hierarchy to infinity, or by the use of singly-conforming tests. Moreover, other forms of t_0 could have been arrived at by modifying some other row of L than the last row.

IX. THE DEFINITION AND MEASUREMENT OF *g*

All the considerations of the preceding sections make it clear that there is not one *g*, but many: *Any three tests* define a *g*, the *g* which would be arrived at by extending these three tests in a hierarchical series to infinity. Starting with these same three tests there are many different hierarchies which can be built upon them, but these all lead ultimately to the same *g*. If we fix our attention on one of the possible infinite series of hierarchical tests which can be built upon these three tests, then any one test of this hierarchy has alongside of it, as it were, a number of other tests any one of which could take its place in the hierarchy but no one of which could coexist with it in the hierarchy: A family of singly-conforming tests.

Any three tests *define* a *g*. But they do not *measure* it exactly unless it happens that the product of two of their correlations exactly equals the third correlation. A fourth test however can always be added (in theory) which will enable the *g* defined by the first three to be measured exactly. This test must *not* be hierarchical with the first three (the defining three). Instead, two of the tetrad-differences must be different from zero and alike, the third tetrad-difference consequently being zero. This however is not sufficient.

Let us call the test which has been added No. 4, and let us number the original three defining tests so that the tetrad-difference which is zero is

$$r_{13}r_{24} - r_{14}r_{23} = 0 \quad (50)$$

Then in order that these four tests may measure exactly that g which tests 1, 2, and 3 define, it is necessary that the fourth test be such that

where

$$\left.\begin{aligned} r_{34} &= l_3 l_4 + m_3 m_4 \\ l_3^2 &= \frac{r_{13} r_{23}}{r_{12}} = 1 - m_3^2 \\ l_4^2 &= \frac{r_{14} r_{24}}{r_{12}} = 1 - m_4^2 \end{aligned}\right\} \tag{51}$$

in which case a man's g can be found exactly and without any indeterminate quantity i from the formula

$$g = \frac{m_4 z_3 - m_3 z_4}{m_4 l_3 - m_3 l_4} \tag{52}$$

where z_3 and z_4 are his standardized scores in tests 3 and 4. All of this is only a special (and the simplest) case of the theory developed in Section V preceding.

It may be well to remind the reader that when it is said here that a certain defined g can thus be measured exactly, what is meant is that the inexactitude due to using a hierarchical team of tests has been abolished. It is assumed however that there is no population-sampling error, the correlations having been found from a population-wide survey: And it is assumed that there is no variation, either real or apparent, in this man's test ability from day to day, no experimental error. Population-sampling error and experimental error will of course in practice come in. But the indeterminate quantity i (which depends on neither of them) has been abolished.

It is necessary to note carefully the proper numbering of the three defining tests as explained above. Note also that only the scores in tests 3 and 4 of this man whose g is being measured are required for equation (52). The original defining tests 1, 2, and 3 define the same g as the trio 1, 2, and 4, but not the same as is defined by 1, 3, and 4 or 2, 3, and 4 if we had started with them.

There are then many g's. Whether we care to call one of them general intelligence is a psychological, not a mathematical question. In practice the psychologist would no doubt hesitate to record his vote if he only knew the original three defining tests: He would wish to see a large number of the tests which formed a hierarchy with those three. In theory, two hierarchies which have three tests in common define the same g. In practice, population-sampling and experimental

error will make this uncertain. Two hierarchies which have only two tests in common need not by any means define the same *g* and will not usually do so.

Finally, the case of three tests which are such that the product of two of their correlations is greater than the third correlation should be mentioned. They also, mathematically, define a *g*. But it is a *g* whose supposition involves the existence of specific factors with unreal weights, weights containing the root of minus one. This *g* no one would suppose to correspond to anything real in the psychological or physiological world. It is only a mathematical expression. And so also in my opinion is *any g*: There is no reason to suppose that *g* suddenly becomes something actual when the product of any two correlations of the defining tests is equal to the third correlation, or when it is less.

X. SUMMARY

An orthogonal matrix *B* is described, a slab of which enables us readily to calculate innumerable transformations of the measures of general and specific ability (ascribed to a population by means of a team of hierarchical tests) into other alternative measures fulfilling all the conditions, illustrating the indeterminate nature of such factors thus measured. It is then shown how by the addition of "singly-conforming" tests to the hierarchy this indeterminateness can be abolished. Incidentally a matrix *D* is arrived at which turns the hierarchical axes so that any chosen one of them goes out of use while all the tests except one remain "two-factor." Finally it is pointed out that *any* three tests *define* a "*g*," and that by the suitable addition of one more test *not* hierarchical with these three this "*g*" can be exactly *measured*, apart from experimental error or population-sampling error which in practice will come in.

MULTIPLE FACTOR ANALYSIS AND PSYCHOLOGICAL CONCEPTS

S. N. F. CHANT

University of Toronto

The scientific status of any assertion rests upon the observable evidence which supports it, which in turn rests upon a knowledge of the circumstances under which the observations were made. With such an emphasis upon the conditions of observation it is inevitable that the observational methods employed should largely determine the meaning of the assertion. For example, the scientific meaning of time is quite different from the perceptual meaning, for it is formulated in terms of the scientific methods which are employed for measuring time. Any scientifically useful definition of time will be stated in terms of scientific measurements and the mathematical analysis of these. Although such a definition may, on occasion, disagree with the perceptual experience of duration, its sanction is found in the methods of measurement, and the extent to which it enables us to formulate scientific laws which extend our control over natural forces.

In psychology this tendency to formulate concepts in terms of method, while not general, has become increasingly prominent. This is particularly the case since the development of methods of correlation. In 1903 Thorndike had sufficient confidence in the method of correlation to assert that; "the mind is a host of highly particularized and independent faculties."[1] In 1904 Spearman evidenced a similar confidence when he stated that; "all branches of intellectual activity have in common one fundamental function (or group of functions), whereas the remaining specific elements seem in every case to be wholly different from that in all others"; and he advocated a correlational psychology for the purpose of positively determining all psychical tendencies.[2]

Such assertions place a tremendous emphasis upon the method of correlation and such methods as are employed for analysing coefficients of correlation. The question is; how far is this justified in psychology?

That psychology has scarcely as much sanction for basing concepts upon methodological considerations as the physical sciences have is

[1] Thorndike: *Educational Psychology*. 1903, p. 39.

[2] Spearman: "General Intelligence Objectively Determined and Measured." *American Journal of Psychology*, Vol. XV, 1904, p. 202.

apparent because of the relative inadequacy of most of the methods in psychology as compared to the methods in the physical sciences. For example, the method of correlation, about which our discussion centres, is often considered to be closely similar to the experimental method of concomitant variations as employed in physical science. Usually the methods are quite different. In physical science in order to determine concomitant variation, two variables are experimentally isolated and made subject to quantitative control. For example, when determining the relationship between the length of a metal bar and its temperature we can vary the temperature by known amounts and determine directly the concomitant change in length. The result can be expressed as a functional equation. Correlation is usually not analagous to this, for we cannot experimentally isolate the variables and quantitatively alter one in order to measure the concomitant change in the other. Thus, we cannot alter a child's intelligence and measure any concomitant change in his chronological age, although the converse may be made to approximate an instance of concomitant change.

In most instances where correlation is used we are not dealing with quantitative change. In our opinion Spearman is in error when he maintains that; "a correlational coefficient, as is now generally known, consists in a number that indicates how far changes in any variable magnitude are accompanied by corresponding changes in another."[1] Usually there is no change to measure; all we are dealing with are fixed deviations from a mean which are unalterable. There is no variation except in the sense of statistical variation which is only alike in name to that type of variation which implies the progressive alteration of a particular quantity. By correlation we measure the average amount of agreement between the fixed values of two quantities expressed as deviations from the mean in a number of discrete instances. This is quite a different thing from the experimental determination of concomitant variations.

Similarly when we employ the regression equation $x = r\frac{\sigma x}{\sigma y}y$ we are not expressing the extent to which the values of x change for the values of y, but merely the extent to which fixed deviations of X from the mean of X correspond to fixed deviations of Y from the mean of Y. In so far as the variables cannot be isolated and quantitatively controlled

[1] Spearman: *The Abilities of Man.* 1927, p. 56.

we are not measuring concomitant change but merely the agreement between fixed amounts.

The meaning of the obtained coefficient is correspondingly different from the meaning of a quantitative law determined on the basis of the experimental method of concomitant variations. The latter provides a basis of control; the former does not. Because we find a high degree of correlation between the length of right legs and the left, it does not follow that if we shorten a right leg the left will become correspondingly short. On the other hand, a coefficient of linear expansion, derived from the experimentally determined concomitant variations of length and temperature, does mean that to increase the temperature will occasion a corresponding increase in length.

The above discussion is sufficient to indicate the relative inadequacy of the method of correlation as compared to the experimental method of concomitant variations. Further indications of its inadequacy could be disclosed through a discussion of the difficulty of rigorously interpreting the amount of agreement expressed by any coefficient of correlation.

Our principal concern, however, is with the more recent method of multiple factor analysis. So much has been written concerning the difficulties of the method of correlation that we may omit further consideration of it, stating merely that for some purposes it is a decidedly useful method, yet relatively inadequate as a basis for formulating psychological concepts. In multiple factor analysis, to all the difficulties inherent in the method of correlation are added further approximations and assumptions which increase the inadequacy of it as a basis for psychological assertions.

Our discussion will largely concern the method of multiple factor analysis as developed by Thurstone.[1] Far from this being a reflection upon Thurstone's work, it is a recognition of the writer's intimate knowledge of the care and understanding, both psychological and mathematical, which Thurstone brings to this work. The writer in no sense deprecates the method of factor analysis; actually it is of great value. Our purpose is rather to clarify and it is hoped extend its value.

In introducing his method, Thurstone makes this rather surprising psychological assertion at the very outset; "Our first question concerns the number of general, independent and uncorrelated factors that are

[1] Thurstone: "Multiple Factor Analysis." *Psychological Review*, Vol. XXXVIII, 1931.

operative in producing a given table of intercorrelations for any number of variables."[1] This statement is in conformity with Spearman's position regarding specific factors. In spite of Spearman's work this assertion remains contradictory to so much of our psychological knowledge which indicates that mental life is highly integrated, that we are justified in inquiring into the assumptions underlying such a claim. There is ample evidence for the alternative claim that all types of mental activity, far from being unrelated or independent, are inextricably interrelated and interdependent. When Thurstone analyses Strong's data on professional interests he finds the following "independent interest factors"; (1) interest in science, (2) interest in language, (3) interest in people, (4) interest in business. Independent in this instance certainly does not mean that they cannot express themselves in the behavior of the same individual, so to this extent at least they are not psychologically independent. Are they then independent within the individual? Hardly, for surely a man's interest in business may increase his interest in language; he may become interested in language for business purposes, and as his interest in business wains so possibly may his interest in language. Similarly with regard to Thurstone's analysis of Moore's data on psychotic symptoms. To stress the independence of the mental disease entities obtained, is to contradict the experience of psychiatrists who ordinarily find the symptoms so interrelated and interdependent that diagnosis is extremely difficult. There is ample evidence against the independence of such factors to justify a careful consideration of the origin of the assertion.

Can we at least admit that these mental factors are uncorrelated? There is more justification for this, because an insignificant coefficient of correlation may not mean that the factors involved are unrelated and independent. A low coefficient of correlation may be obtained between two highly interdependent factors, but due to inadequacies in measurement, and the effect of uncontrolled variables, the coefficient may be insignificant, and strictly speaking uncorrelated only means having a low coefficient of correlation. Spearman's emphasis upon attenuation acknowledges the effect of such variables, and his objection to Thorndike's claim for independent mental factors is based upon the fact that inaccuracies of measurement occasion low coefficients.

---

[1] Thurstone: "Multiple Factor Analysis." *Psychological Review*, Vol. XXXVIII, 1931, p. 406.

Why then does Thurstone make such an assertion at the very outset? The first indication of why he does so appears to be in order to justify a certain step in his method of analysis. On page 409[1] he states; "$\frac{\Sigma s_a s_b}{N} = a_1 b_1 \frac{\Sigma x_1^2}{N} + a_2 b_2 \frac{\Sigma x_2^2}{N} + a_3 b_3 \frac{\Sigma x_3^2}{N}$ in which the cross products vanish because x_1, x_2, x_3 are uncorrelated." But dropping these cross products needs no such justification. Frankly, we are unable to take account of them so we drop them. This is an approximation but no harm results in so far as we can estimate the amount of distortion occasioned in our data by such an approximation. As Thurstone points out this can be done by comparing the experimentally obtained coefficients with the coefficients obtained by equations of the following type; $r_{ab} = a_1 b_1 + a_2 b_2 + a_3 b_3$. It is quite unnecessary therefore to force psychological meanings in order to justify this approximation. Anyone who claims that such an approximation vitiates the whole method of analysis is so lacking in mathematical perspective as to be ignored.

Thurstone's emphasis upon independent uncorrelated mental factors appears to be made in order to meet the demands of the theory of errors rather than the facts of psychology. But when dealing with coefficients of correlation, to invoke the theory of errors seems largely a rationalization to justify certain approximations which may as well be frankly admitted. As will be indicated later such insistence upon the independence of the factors makes Thurstone's method very difficult to explain, and there is less possibility of contradiction in the method if we accept the approximations and do not attempt to justify them on the basis of a psychologically untenable assertion.

Indeed the whole usage of the term factor must come in for careful consideration, for it is becoming much more prevalent in psychology with the development of these methods of factor analysis. Certainly the method of analysis involves mathematical factors. But are there similar mental factors? One suspects that both Spearman and Thurstone in the use of the term factor have fallen into the error of "having fashioned for ourselves such a name as, say, judgment, then—because the name remains the same—we fall into the belief that we are treating of an entity always the same."[2]

---

[1] Thurstone: "Multiple Factor Analysis." *Psychological Review*, Vol. XXXVIII, 1931, p. 409.

[2] Spearman: *The Abilities of Man.* 1927, p. 39.

Accepting Thurstone's claim that Spearman's two factor analysis is a special case of his own more general analysis, our remarks will be confined to Thurstone's method. Omitting the geometrical formulation of the proof,[1] Thurstone's method does what any satisfactory statistical formulation should do, and no more. It systematizes a body of data so that certain interpretations are made possible; but it does not supply these interpretations. Thurstone indicates this when he states; "By inspection it is possible to name the first factor although the statistical procedures do not of course concern these matters of describing or naming the factors."[2] Similar to other types of statistical analysis, Thurstone's method is a special type of classificatory technique. Its value lies in the organization of data, rather than in the explanation of these data. Reduced to a simple statement the method classifies a set of intercorrelations and on this basis makes possible the description of certain general categories or types of behavior. The final validation of the classification lies, as Thurstone states, in the extent to which the classification is psycholgocially intelligible.

The first step in this method of classification is that of differentiating a certain group of coefficients in a table of intercorrelations from all other coefficients in the table. This is the initial dichotomy that is required for almost any system of classification. Thurstone does this by selecting a particular test, *a*, and grouping all tests which correlate positively with *a*. This group of tests is designated *s*. Test *a* is chosen as the one having the highest sum of intercorrelations with all the other tests in the table of intercorrelation. It therefore provides a more comprehensive group of tests correlating positively with it, than would any other test. The tests which fall in group *s* therefore constitute a comprehensive group of tests differentiated from all others in the table of intercorrelations.

The relative degree of internal coherence within group *s* can be indicated by the average amount of intercorrelation found in the group. The extent to which any test in the whole table of intercorrelations fits into group *s* can be expressed as an index of the sum of its correlations

[1] Indeed the geometrical formulation offers a difficulty in that it appears to be contradictory to the fact that when all the coefficients are positive, the method yields, like Spearman's method only one general factor. But geometrically, coefficients of negative sign should not be necessary in order to define a *y* axis in addition to an *x* axis.

[2] Thurstone: "Multiple Factor Analysis." *Psychological Review*, Vol. XXXVIII, 1931, p. 426.

with the tests in group s, to the sum of all the intercorrelations within group s. This is given for each of the tests in the table of intercorrelations, by such equations as the following;

$$a_1 = \frac{\Sigma r_{as}}{\sqrt{\Sigma r_{ss}}}$$

$$b_1 = \frac{\Sigma r_{bs}}{\sqrt{\Sigma r_{ss}}}$$

$$\vdots$$

$$z_1 = \frac{\Sigma r_{zs}}{\sqrt{\Sigma r_{ss}}}$$

Group s having served its purpose by providing the above indices plays no further part in the method. On the basis of the indices we may describe the first general category or type of behaviour measured. The forms of behaviour measured by those tests with high indices will largely determine the definition of the general category, those with low indices will not play much part in the definition since they do not fit into the general category or type to any appreciable extent.

In so far as this first general class or category does not account for all the intercorrelations in the table, further categories must be found. If all the coefficients are positive this method yields only one general category in as much as Σr_{as} will equal Σr_{ak} where Σr_{ak} is the sum of all the intercorrelations of test a with all the tests in the whole table of intercorrelations; this will be the case for each test. In other words group s will include all the tests of the table. This would be Spearman's special case.

In so far as this is not the case additional general categories of behaviour are required to exhaust the possibilities of the table of intercorrelations. In order to obtain a second category we require a criterion to employ as test a was employed to obtain category one. In order that the second category may be dissimilar from the first, and thus facilitate classification, we select some test as its basis which is dissimilar to test a. To do this we list all the tests having a coefficient of correlation with test a of between $+.5$ and $-.5$, and select from these that test which has the highest sum of intercorrelation with all the tests in the whole table of intercorrelations. Then by selecting all the tests which correlate positively with this test, we obtain a

comprehensive group of tests which is relatively dissimilar to the tests in group s. This group of tests is termed group t.

We can now determine, just as with regard to group s, the weight of each of the tests in the whole table of intercorrelations with regard to group t. Then on the basis of those tests which are similar in that they are heavily weighted we can describe a second general category or type of behaviour. The equations for these weights will be as follows:

$$a_2 = \frac{\Sigma r_{at} - a_1 \Sigma t_1}{\Sigma t_2}$$

$$b_2 = \frac{\Sigma r_{bt} - b_1 \Sigma t_1}{\Sigma t_2}$$

$$\vdots$$

$$z_2 = \frac{\Sigma r_{zt} - z_1 \Sigma t_1}{\Sigma t_2}$$

These weights express the extent to which any test fits into the set of tests t, less the extent to which this is occasioned by its weight with regard to category one. This is so because $\Sigma r_{at} = a_1 \Sigma t_1 + a_2 \Sigma t_2$ and by subtracting $a_1 \Sigma t_1$ we find Σr_{at} freed from the influence of the general type of behaviour defined by category one. Similarly Σt_2 is the internal correlation in group t, less the extent to which this is occasioned by the influence of the type of behaviour described by category one, since $\Sigma r_{tt} = (\Sigma t_1)^2 + (\Sigma t_2)^2$. The equations for determining the weights with regard to general category two are therefore identical in form to those for determining the weights with respect to category one, except that in the latter case we subtract from them any influence due to the tests being concerned with the type of behaviour described by category one. This removes any influence of the first general type of behaviour in the description of the second general type. In this sense the description of type two is independent of type one. But the description of type one is not similarly independent of type two, because the subtraction takes place after type one has been defined. Hence it is difficult to see just how type one can be considered as separate from and independent of type two.

The equations for determining weights with regard to a third general type of behaviour are identical to those for determining the weights for types one and two; except that in the case of type three we subtract any influence which the weights with regard to types one and

two may have upon the weights for type three. In this manner type three may be described as independent of types one and two. Again, however the subtraction takes place after we have determined types one and two so that these are not freed from the effect of general type of behaviour three. Thus categories one and two cannot be said to be independent of category three.

Similar weights may be assigned with regard to additional general categories until the method is "carried out far enough to lock the coefficients with discrepancies somewhat greater than the chance errors of the given coefficients."[1]

That Thurstone recognizes the method as essentially one of classification is shown by his statement that an estimate of the number of factors that will be needed to account for a table of correlations can be made upon steps six, eleven, sixteen, twenty-one, twenty-six, thirty-one, and so on. That these steps are wholly classificatory can be seen at once by referring to his outline of steps for the procedure.

Our whole discussion indicates that this method of multiple factor analysis can scarcely be said to yield independent mental factors. Does this then invalidate the method? Far from doing so it makes the result of such analysis more psychologically intelligible and hence more valid. That which results from the analysis is not a number of independent mental factors but a classification of behaviour into certain general descriptive categories or types. The extent to which any test measures or describes any of these types of behaviour is indicated by its weight with regard to that type. Hence such a method of analysis does not dictate our psychological concepts, but only systematizes our data so that a classification of behaviour into broad general types becomes possible. This is just as valuable in every respect as the postulation of independent uncorrelated mental factors, and it has the additional advantage of being psychologically intelligible. Thus with respect to Thurstone's analysis of Strong's data on professional interests, we find four general types of interest; interest in science, interest in language, interest in people, and interest in business; just as fear and anger are two general types of emotional behaviour. Also we may describe certain professions as exhibiting these interests in varying degrees, as well as describe the interests of a person with regard to them.

---

[1] Thurstone: "Multiple Factor Analysis." *Psychological Review*, Vol. XXXVIII, 1931, p. 421.

To accept the method of multiple factor analysis as a method of classification overcomes an additional difficulty. If new tests are introduced into the table of intercorrelations the weights with reference to the categories become altered. If the analysis yielded independent uncorrelated mental factors these should retain their integrity in spite of the additional tests. There is a certain finality about the theory of mental factors which is not present in a mere classification of behaviour into general types. Any system of classification can be altered according to the purpose for which it is made and according to the scope required. Thus we are free to choose that form of classification which best suits the facts, rather than attempt to force the facts to conform to the system of classification. That which is required in science is a body of rigidly determined data and a flexible method for analysing these, rather than the converse.

Each of the types of behaviour may be defined as an "ability" if this is desired, just as readily as a "factor" can be so defined. Indeed all the evidence we have for so-called abilities is that there are observable types of behaviour which persons exhibit in varying degrees. These varying degrees merely express the extent to which a person's behaviour can be described in terms of the type. Indeed a classification of behaviour into general types supplies all we ever know and all we ever need to know concerning abilities.

Summing up in conclusion we may say that the method of multiple factor analysis provides an extremely useful method for classifying quantitative data. On the basis of the classification certain general types of behaviour may be described. When we go further and try to make our psychological concepts conform to the theory of errors because we have made certain mathematical assumptions in the working out of the method, we add nothing to our psychological knowledge and tend to discredit an otherwise very valuable technique.

PERSONALITY TRAITS BY FIAT. I. THE ANALYSIS OF THE TOTAL TRAIT SCORES AND KEYS OF THE BERNREUTER PERSONALITY INVENTORY[1,2]

IRVING LORGE

Institute of Educational Research, Teachers College, Columbia University

Bernreuter's[3] basic contribution to measurement of personality is that of regarding the subject's response to each item of a test as an indicator of several types of adjustment, of several modes of reaction or of several traits. Bernreuter was the first person to apply this principle of multiple scoring to personality, although both Cowdery and Strong have made use of the principle in the measurement of "interests."

Bernreuter, after an extensive study of personality tests existing at the time of his compilation, selected one hundred twenty-five items purporting to measure aspects of personality adjustment from four scales: 1. Thurstone and Thurstone's *Neurotic Inventory*;[4] 2. G. W. Allport's[5] *A-S Test of Ascendance-Submission;* 3. Laird's[6] *C2 Test of Introversion and Extroversion;* and 4. Bernreuter's[7] *S-S Test of Self Sufficiency.* Each item was framed as a question to which the subject could respond by encircling "Yes," "No," or "?" as he agreed, disagreed, or was in doubt about the question. Each response to each item was given a scoring weight based upon the diagnostic value of the response in distinguishing between low- and high-scoring groups in each of the four scales from which the items were culled. This analysis yielded four scoring keys in which each subject's response was

[1] Acknowledgement is hereby made of the services rendered by the personnel furnished by the Works Division, Emergency Relief Bureau of New York City on Project 89FB-125X.

[2] This study is part of a larger study in Interests, Attitudes, and Motives supported in part by a grant from the Columbia University Council for Research in the Social Sciences.

[3] Bernreuter, R. G.: *The Personality Inventory.* Stanford Univ. Press, 1931.

[4] Thurstone, L. L. and Thurstone, T. G.: "A neurotic inventory." *J. Soc. Psychol.*, Vol. I, 1930, pp. 3–30.

[5] Allport, G. W.: "A test of ascendance-submission." *J. Abn. and Soc. Psychol.*, Vol. XXIII, 1928, pp. 118–136.

[6] Laird, S. A.: "Detecting abnormal behavior." *J. Abnormal and Soc. Psychol.*, Vol. XX, 1925, pp. 128–141.

[7] Bernreuter, R. G.: "The measurement of self sufficiency." *J. Abn. and Soc. Psychol.*, Vol. XXVIII, 1933, pp. 291–300.

given a differential weight in four "traits." The four traits so keyed were: 1. B1-N, neurotic tendency; 2. B2-S, self sufficiency; 3. B3-I, introversion-extroversion; and 4. B4-D, dominance-submission.

TABLE I.—THE INTERCORRELATIONS AMONG THE TOTAL TRAIT SCORES, AND THE INTERCORRELATIONS AMONG THE WEIGHTS FOR THE TRAIT KEYS BY "YES," "NO," "?" FOR THE BERNREUTER PERSONALITY INVENTORY
Where 1 = B1-N, neurotic tendency; 2 = B2-S, self sufficiency; 3 = B3-I, introversion-extroversion; 4 = B4-D, dominance submission

	Bernreuter total scores				Bernreuter weights for yes responses	Bernreuter weights for no responses	Bernreuter weights for ? responses
	Lorge	Bernreuter		Flanagan			
	$n = 103$	$n = 70$	$n = 128$	$n = 305$			
r_{12}	−.4739	−.39	−.39	−.442	−.2653	−.2535	−.2201
r_{13}	.8932	.95	.93	.978	.8043	.8035	.2566
r_{14}	−.7972	−.87	−.82	−.797	−.7439	−.6956	−.1759
r_{23}	−.3766	−.28	−.28	−.362	−.0824	−.0921	−.0572
r_{24}	.6277	.55	.50	.571	.4337	.4278	.2149
r_{34}	−.6692	−.75	−.73	−.694	−.5341	−.4894	−.1000

TABLE II.—THE MEANS AND STANDARD DEVIATIONS OF THE TOTAL TRAIT SCORES AND OF THE TRAIT WEIGHTS BY "YES" "NO" AND "?" FOR THE BERNREUTER PERSONALITY INVENTORY
Where 1 = B1-N, neurotic tendency; 2 = B2-S, self sufficiency; 3 = B3-I, introversion-extroversion; 4 = B4-D, dominance-submission

	Bernreuter total scores (Lorge data)	Bernreuter weights for yes responses	Bernreuter weights for no responses	Bernreuter weights for ? responses
	Means			
1	−66.0582	1.3040	−0.9600	.6480
2	43.1263	−0.3280	0.3600	.2080
3	−33.7378	1.0640	−0.7760	.2720
4	55.5437	−0.2080	0.4960	.8320
	Standard Deviations			
1	71.4036	2.4987	2.5500	1.2915
2	50.2833	2.2263	2.2249	1.3464
3	45.1642	1.5427	1.6043	1.2546
4	54.2719	2.0834	2.1490	1.2821

The Bernreuter Personality Inventory because of its wide usage in research, in vocational guidance, and in clinical studies, was administered to a group of one hundred three adults regularly employed as

subjects in an adult learning project. These adults were of both sexes ranging in age from twenty to seventy, and ranging in intellect from CAVD three hundred sixty-nine to CAVD four hundred forty (on Army Alpha the range was from thirty-eight to two hundred seven).

For each person, each of the four trait scores was computed. The mean total trait scores and the standard deviations are reported in Table II. The intercorrelations among the four total trait scores were calculated. These intercorrelations are reported in Table I, together with intercorrelations among the total trait scores reported by Bernreuter in his manual, and those reported by Flanagan[8] in his recent factor analysis of personality. Comparison of the correlations indicates substantial agreement among the four sets of intercorrelations from four different groups. The data further show that none of the Bernreuter total trait scores are independent of each other. Bernreuter[9] has already shown that trait B1-N (neurotic tendency) and trait B3-I (introversion-extroversion) might be regarded as alternative measurements of the same phenomena. There can be little doubt of this in view of correlations of the order of .89, .95, .93, and .98. Not only is there considerable community of traits between B1-N and B3-I, but inspection of all orders of partial correlation coefficients indicates other community in the remaining two traits.

These partial correlations are reported in Table III.

The correlation between neurotic tendency and introversion-extroversion partialling self-sufficiency and dominance-submission is +.80. The +.80 is indicative of a high degree of community. The correlation between self-sufficiency and dominance-submission partialling neurotic tendency and introversion-extraversion is +.46. Bernreuter had indicated that there are three measurable traits in his inventory. The partial correlations, however, point the fact that the traits as measured by the Bernreuter are not sufficiently independent of each other for the clinical psychologist to view the raw scores as discrete trait measurements. As a further demonstration of the lack of discreteness among the trait scores, the intercorrelation between the sum of B1-N (neurotic tendency) and B3-I (introversion-extroversion) and the sum of B2-S (self-sufficiency) and B4-D (dominance-submission) was computed. The correlation is −.68. If the measure

[8] Flanagan, J. C.: *Factor Analysis in the Study of Personality.* Stanford University Press, 1935.

[9] Bernreuter, R. G.: "The imbrication of tests of introversion-extroversion and neurotic tendency." *J. Soc. Psych.*, Vol. V, 1934, pp. 184–201.

called neurotic tendency, and that called introversion-extroversion are considered as two approximations of the same trait, the Spearman-Brown reliability of the sum of the two scores may be estimated at +.94. Similarly if self-sufficiency and dominance-submission are

TABLE III.—THE INTERCORRELATIONS BETWEEN PAIRS OF TOTAL TRAIT SCORES PARTIALING OUT THE INFLUENCE OF THE OTHER TRAIT SCORES, AND THE SIMILAR PARTIALS AMONG THE INTERCORRELATIONS OF WEIGHTS FOR THE TRAIT KEYS BY "YES," "NO," AND "?" FOR THE BERNREUTER PERSONALITY INVENTORIES

Where 1 = B1-N neurotic tendency; 2 = B2-S self sufficiency; 3 = B3-I introversion-extroversion; 4 = B4-D dominance-submission

	Bernreuter total scores (Lorge data)	Bernreuter weights for		
		Yes responses	No responses	? Responses
$r_{12.3}$	−.3301	−.3360	−.3028	−.2129
$r_{12.4}$	.0563	.0952	.0679	−.2683
$r_{13.2}$	.8761	.8142	.8099	.2300
$r_{13.4}$	.8018	.7204	.7392	.2351
$r_{14.2}$	−.7280	−.7238	−.6716	−.1350
$r_{14.3}$	−.5960	−.6257	−.5844	−.1412
$r_{23.1}$	.1170	.2287	.1938	−.0007
$r_{23.4}$	.0751	.1950	.1870	−.0236
$r_{24.1}$	.4701	.3667	.3619	.1835
$r_{24.3}$	.5458	.4626	.4407	.2087
$r_{34.1}$	.1580	.1617	.1711	−.1214
$r_{34.2}$	−.5994	−.5550	−.4999	−.1521
$r_{12.34}$	−.0067	−.0675	−.1063	−.2818
$r_{13.24}$	.8011	.7180	.7413	.2376
$r_{14.23}$	−.5269	−.5627	−.5271	−.1942
$r_{23.14}$	.0503	.1844	.2036	−.0926
$r_{24.13}$	.4606	.3432	.3410	.1846
$r_{34.12}$	.1162	.0856	.1014	−.1900

regarded as two approximations of the same trait, then the estimate of the reliability of the sum is +.77. The correlation coefficient between the sum of N + I with the sum of D + S may then be corrected for attenuation. The resulting corrected correlation coefficient is then −.81.

Self-sufficiency plus dominance-submission might be regarded as the counterpart of neurotic tendency plus introversion-extroversion. Flanagan subjected the total trait score intercorrelations obtained from a population of three hundred five eleventh grade boys to a Hotelling[10] principal component analysis. As a result he finds two factors: The first factor, representing about seventy-eight per cent of the non-chance variance, has large positive weights for Neurotic Tendency and Introversion, a rather large weight of opposite sign for Dominance, and a somewhat smaller negative weight for Self-Sufficiency. . . . The second factor, representing about eighteen per cent of the non-chance variance, is seen to have a large positive weight for Self-Sufficiency, a smaller positive weight for Introversion-Extroversion, and slight positive weights for Neurotic Tendency and Dominance." From Flanagan's study it is apparent that the four traits can easily be reduced to two.[11]

By either the Flanagan or the Lorge analysis, it is apparent that less than four, and probably not more than two independent traits can be used to account for the "traits" of the Bernreuter.

If the keys for the scoring of each trait are independent of each other, then the relationships so far reported must be considered a function of the organization of personality in the group of subjects tested. Such an organization of personality might indeed lead to the discovery of important research values for the inventory. Each item in the Bernreuter is marked by encircling "Yes" or "No" or "?." The scoring of the inventory provides differential weighting for each response for each of the four traits. There are four sets of keys corresponding to each trait in which each response "Yes," "No" or "?" is given a weight from +7 to −7. The intercorrelations of the weights given "Yes" responses in each of the four traits for the 125 items were computed; so were the corresponding intercorrelations for the "No" responses and for the "?" responses. The data are reported in Table I. The intercorrelations for the "Yes" and "No" keys are strikingly similar to the intercorrelations among the total trait scores. The intercorrelations for the "?" keys are in the same direction as for the trait scores. It is apparent that the interrelations existing among the total trait scores exist among the keys. As a further dem-

[10] Hotelling, H.: "Analysis of a complex of statistical variables into principal components." *J. Educ. Psychol.*, Vol. XXIV, 1933, pp. 417–441 and 498–520.

[11] Flanagan used intercorrelations which were corrected to minimize the effect of the correlation between errors among trait measurements.

onstration, the partial correlations were also computed. They show the same direction and almost the same magnitude as was found in the total scores. The relationship in the keys indicates that the correlations in the scores are attributable to the weighting system. The differences between the magnitude of the correlations in the keys, and that in the scores is attributable to the mixing of the "Yes," "No," and "?" keys in determining trait scores. Karl Pearson[12] has demonstrated that correlation is created or affected spuriously when a mixture of heterogeneous materials is made.

The facts reported in Tables I and III demonstrate the pitfalls in assigning weight scores on the basis of diagnostic value of an item in distinguishing a low- and a high-scoring criterion group without partialling out the effect of other traits. The partial correlation might have been of value in developing independent keys. Flanagan, however, developed two trait keys independent of each other to measure the two factors he isolated through Hotelling's analysis. The independence of Flanagan's keys for factors one and two is shown by the correlations of +.0422 for the "Yes" keys, of −.0028 for the "No" keys, and of −.0376 for the "?" keys. The Flanagan keys are based upon corrected correlation coefficients among the total trait scores.

Flanagan has tentatively named his two new traits Self-confidence and sociability. He realizes fully well that naming the factor does not make it a trait. He, nevertheless, is approaching the problem from a healthy point of view.

Personality traits cannot be created by the psychologist. If the concept of personality is to have meaning, it must be conceived as an aspect of the individual—an aspect susceptible of quantification. Naming a trait does not make it a trait. To be useful in psychology, a trait must be regarded as freed from the influences of other traits, and other aspects of the individual. The separation of neurotic tendency from introversion-extroversion is a desirable goal, but it is a goal not easily achieved by consideration of measurements that correlate too highly. Purity of trait, from the point of view of consistency of measurement, from the point of view of reliability of measurement, and from the point of view of independence of measurement, is the only hope in the all too elusive field of personality diagnosis. Personality traits cannot exist by fiat alone.

---

[12] Pearson, K., Lee, A., and Bramley-Moore: *Mathematical contributions to the theory of evolution.* VI. 192A, 1898, pp. 257–330.

CHANCE AND PERSONALITY INVENTORY SCORES

ROBERT G. BERNREUTER

The Pennsylvania State College

In an earlier paper,[3] it was pointed out that the distributions of scores customarily obtained on any of the scales of the Personality Inventory were very different from those that would be expected were chance alone operating. In that paper, the evidence consisted of a comparison of the standard errors of single scores with the standard deviations of the total distributions.

The present paper is concerned with distributions which are obtained when chance alone determines the answers to the questions. They are compared with the distributions obtained from a group of four hundred twenty-seven male college students, selected because this group seemed typical of the results usually obtained from groups of unselected individuals.

METHOD OF OBTAINING CHANCE SCORES

Two hundred scores were first obtained on the B1-N scale by drawing one hundred twenty-five pennies at random two hundred times from a box in which they had been well shaken. If the first penny drawn were a head, the first question was scored as though it had been answered by "yes." If the first drawn were a tail, it was scored as though answered by "no." The same procedure was followed until each of the one hundred twenty-five questions had been scored. Totals were taken for first sixty-three items as well as for all the items in order to compute the half vs. half reliability of the chance scores. The B2-S scale was then scored by the same procedure, followed by each of the other two scales. Separate drawings were made for each. The question marks were not considered, because of the relative infrequency with which they are used by unselected subjects. The inclusion of the question marks to the extent to which they are used by such subjects would probably have modified the obtained distributions very slightly, but would have considerably increased the difficulty of collecting the data.

THE MEAN CHANCE SCORES, AND THEIR SIGNIFICANCE

In Table I, column 1, are shown the mean scores obtained for each of the scales. A comparison of these means with those obtained from

four hundred twenty-seven male college students,* given in column 2, shows that the means of the chance scores are very different from those obtained from college males. In column 4 of this table are given the critical ratios of the differences between the chance means and the college male means. These critical ratios were found by dividing the

TABLE I.—THE MEANS OF THE CHANCE SCORES, AND THEIR SIGNIFICANCE

	(1) Chance mean	(2) College male mean	(3) Sum of mean weights	(4) Critical ratio of differences (1) − (2)	(5) Critical ratio of differences (1) − (3)	(6) Critical ratio of differences (2) − (3)
B1-N.....	22.30	−52.95	21.50	16.03	.36	18.06
B2-S.....	4.25	24.91	3.00	5.41	.77	8.39
B3-I.....	20.00	−27.36	17.50	16.50	1.77	17.96
B4-D.....	21.00	46.32	18.00	7.06	2.01	8.69

differences by the standard errors of the differences. They vary from 5.41 to 16.50, and clearly indicate that the differences are very significant. In column 3 of this table is given the algebraic sum of the means of the "yes" and "no" weights for each scale. These values approximate the mean chance scores. The critical ratios of the differences, as given in column 5, show that in no instance does the mean chance score vary significantly from the sum of the means of the weights. The critical ratios presented in column 6 indicate that the mean scores obtained from college males are significantly different from the sum of the means of the weights.

These differences must be due to the fact that certain of the answers are chosen more frequently than others. The college males tended to choose answers that carried negative weights on the B1-N and B3-I scales, and positive weights on the B2-S and B4-D. The causes for the preference of certain answers over others is, of course, not disclosed by the present study, but that such a preference exists is clearly demonstrated. If chance alone operated to determine the answers of college males, very different mean scores would be obtained for them.

* The data concerning the four hundred twenty-seven male college students are taken from [1].

THE STANDARD DEVIATION OF THE CHANCE SCORES, AND THEIR SIGNIFICANCE

In Table II, column 1, are presented the standard deviations of the distributions of chance scores. In column 2 of this table, the standard deviations obtained from the four hundred twenty-seven college males are shown. The critical ratios shown in column 4 vary from 14.18 to 18.25, and indicate that the dispersion of chance scores is

TABLE II.—THE STANDARD DEVIATIONS OF THE CHANCE SCORES, AND THEIR SIGNIFICANCE

	(1) Chance SD	(2) College male SD	(3) College male SD of single score	(4) Critical ratios of differences (1) − (2)	(5) Critical ratios of differences (1) − (3)
B1-N	31.78	85.17	26.99	16.08	2.61
B2-S	23.10	53.99	21.90	14.18	.87
B3-I	20.75	51.62	18.09	15.07	2.21
B4-D	21.10	67.36	22.96	18.25	−1.41

significantly different from the dispersion of college males' scores. In column 3 are given the standard errors of single scores* as computed from the college male data. These values approximate the standard deviations of chance scores; the critical ratios in column 5 indicate that they do not differ significantly from them.

The very great differences in the dispersion of chance scores and of college males' scores indicate that factors other than chance operate to determine the distribution of scores of college males. The dispersion of college males' scores is greater than would be expected from chance. This could result only if the college males tended to choose consistently answers that carried weights in a given direction, positive or negative. Interpreting this is terms of the B1-N scale, the dispersion of scores obtained from college males results from the consistent choice by them of answers that indicate emotional stability, or, on the other hand, emotional instability. This consistent choice of one type of answer is

* As computed by the formula

$$\sigma_{1 \cdot \infty} = \sigma\sqrt{1 - r_{11}}$$

In Table V,[2] the *probable errors* of single scores were erroneously labeled standard errors of single scores.

what is referred to when the term "trait" is used. The differences between the chance and the college male distributions of scores justifies the use of the concept of "traits."

The values given in Table II, column 3, indicate that the measurement of the traits by the various scales is subject to substantial error. The values in column 5 indicate that this error approximates in size the dispersion in scores due to chance factors.

CHANCE AND RELIABILITY

The coefficients of reliability obtained for the chance scores are presented in Table III. The half-vs.-half coefficients are so small in

TABLE III.—THE RELIABILITY OF CHANCE SCORES

	Chance scores		College males
	Half versus half	Corrected*	
B1-N	.12 ± .07	.22	.90
B2-S	−.06 ± .07	−.11	.84
B3-I	−.03 ± .07	−.06	.88
B4-D	−.17 ± .07	−.30	.88

* By the Spearman-Brown formula.

comparison with their standard errors that the use of the Spearman-Brown prophecy formula might well be questioned. It is apparent that the scores obtained from chance approximate zero in reliability. It is also apparent that the relatively high reliabilities obtained from college students can not be the result of chance alone. They, too, must be due to the existence of factors which cause individuals to choose consistently answers which are weighted in a given direction. Furthermore, these high reliabilities are obtained only because, in the construction of the scales, weights were assigned to answers in a manner consonant with the factors motivating the activities (the so-called "traits") of the college students.

CONCLUSIONS

The comparison of the means, standard deviations and reliabilities of scores obtained on the Personality Inventory by chance with those obtained from a representative group of humans makes it apparent that factors other than chance operate to determine which answers will be given to the various questions. These additional factors,

undisclosed in the present study, cause the subjects to choose answers which are weighted in a single direction more consistently than does chance alone. The existence of these factors justifies the use of the concept of "traits." The relatively high reliability of the scales demonstrates that in the construction of the scales, weights were assigned to the answers in accordance with the traits existing in the human subjects. The error in the measurement of the various traits by the various scales approximates the dispersion of scores on the scales that is due to chance.

BIBLIOGRAPHY

1. Bernreuter, R. G.: *The Evaluation of a Proposed New Method for Constructing Personality Trait Tests.* Unpublished Ph.D. dissertation, Stanford University, 1931.
2. Bernreuter, R. G.: "The Theory and Construction of the Personality Inventory." *J. Social Psychol.*, Vol. IV, 1933, pp. 387–405.
3. Bernreuter, R. G.: "The Validity of the Personality Inventory." *Personnel Journal*, Vol. XI, 1933, pp. 383–386.

SOME OBSERVATIONS ON MULTIPLE-FACTOR ANALYSIS

OLIVE R. RUSSELL

Ontario College of Education, University of Toronto

In a recent study involving factor analysis, the writer found that when the correlation matrixes were analysed by the Thurstone method of "Multiple Factor Analysis"[1] and by the Spearman tetrad technique[2] the results were somewhat at variance. This variance was conspicuous both for a large table of intercorrelations (twenty-three variables) and for several smaller matrixes (six variables each). It was particularly striking where, according to the tetrad technique, group factors over and above a common factor, were indicated. These findings led the writer to devise a hypothetical matrix in which group factors were unquestionably operating. Analysis of this matrix (and others) seemed to illuminate the contrast between the two techniques.

The purpose of the original study[3] was to determine what factors it was necessary to postulate in order to account for intercorrelations of scores obtained on "Verbal," "Mathematical," and "Form Perception" tests. The test battery consisted of twenty-three tests (one thousand one hundred fifty-four items); it was administered to nearly four hundred pupils of the first and second years of a Scottish High School (average age thirteen years). The data were analysed separately for each of the following groups of subjects: First Year Boys, First Year Girls, Second Year Boys, Second Year Girls. For one of these four groups of subjects, the intercorrelations for the twenty-three single tests were computed; for all four groups of subjects, the intercorrelations were computed for various combinations of tests. Each of these smaller matrixes consisted of six variables—two "Verbal," two "Mathematical," and two "Form Perception."

---

[1] Thurstone, L. L.: "Multiple Factor Analysis." *Psychol. Rev.*, Vol. XXXVIII, Sept., 1931, pp. 498–520.

[2] Spearman, C.: *The Abilities of Man* (*appendix*). London, Macmillan & Co., Ltd., 1927.

Holzinger, K. J.: *Statistical Resumé of the Spearman Two-Factor Patterns.* Chicago, University Chicago Press, 1930, pp. 43.

[3] Russell, O. R.: *A Study of the Factors Underlying the Intercorrelations of "Verbal," "Mathematical" and "Form Perception" Tests.* As yet unpublished.

Analysis of each of the matrixes by the Thurstone method of Multiple Factor Analysis (published in 1931) gave no indication of the presence of more than one common factor, while the Spearman technique indicated that, for three of the four groups of subjects, group factors as well as one common factor must be postulated.

The following matrix is one computed for the Second Year Boys ($n = 70$).

CORRELATION COEFFICIENTS ($n = 70$)

Variables		V_1	V_2	M_1	M_2	FP_1	FP_2
"Verbal"	V_1	....	.830	.533	.727	.326	.329
	V_2	.830	....	.596	.739	.412	.432
"Mathematical"	M_1	.533	.596	....	.752	.324	.484
	M_2	.727	.739	.752	....	.520	.619
"Form perception" (selected	FP_1	.326	.412	.324	.520	....	.705
tests)	FP_2	.329	.432	.484	.619	.705	....

When this matrix was analysed by the Thurstone method the first factor loadings were found to be as follows:

V_1	V_2	M_1	M_2	FP_1	FP_2
.79	.84	.78	.92	.69	.75

No second factor was indicated by this technique.

When this same matrix was analysed by the Spearman method it was found that there were several significant tetrad differences (14 TD's exceeded 3½ times their PE's). The distribution of tetrad differences, as indicated in the graph, led to the conclusion that it was necessary to postulate group factors in addition to any factor common to all variables.

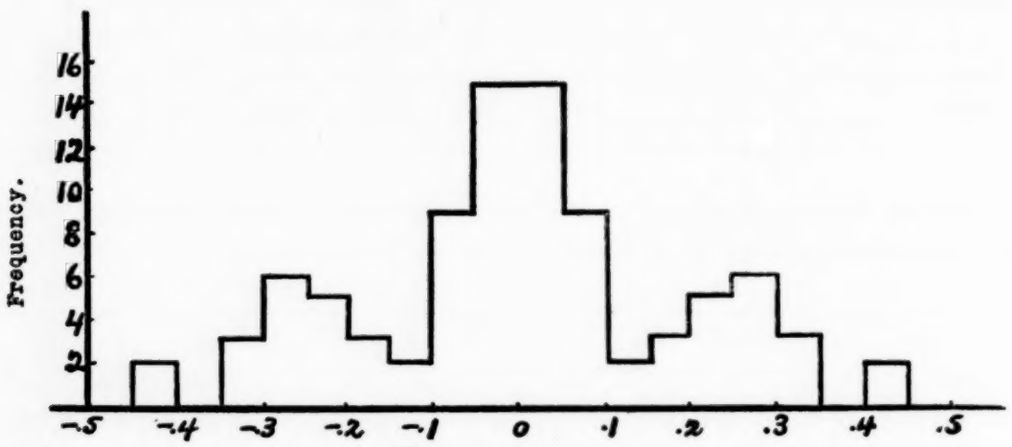
This type of conflict is frequently encountered. It is very evident in the following matrix, which has been constructed artificially, so as to provide, for analysis, a case in which group factors are unquestionably indicated.[1]

[1] The hypothetical matrix has been constructed so as to accentuate the factor-pattern presented by the matrixes of actually obtained correlations, as found in the preceding table.

HYPOTHETICAL MATRIX

Variables	*a*	*b*	*c*	*d*	*e*	*f*
a	..	.9	.5	.5	.1	.1
b	.9	..	.5	.5	.1	.1
c	.5	.5	..	.9	.5	.5
d	.5	.5	.9	..	.5	.5
e	.1	.1	.5	.5	..	.9
f	.1	.1	.5	.5	.9	..

Negative correlations have been omitted purposely from this matrix, both because they are seldom found among the intercorrelations of tests of mental ability and because they were not present in



Distribution of Tetrad Differences

the original experimental data. Let us assume that all the correlations have equal reliabilities, each of which is very high. Since $r_{ab} = .9$ and $r_{ef} = .9$, but each of the correlations of tests *a* and *b* with each of *e* and *f* is only .1, it is evident that there is a group factor operating in tests *a* and *b* that is very different from that operating in tests *e* and *f*. Therefore at least two group factors must be postulated in order to account for the intercorrelations of these tests. In order to account for the whole matrix it seems necessary to postulate three strong group factors, and one very weak factor, common to all six variables.

Analysis of this hypothetical matrix, according to the Thurstone technique (1931), suggests the presence of one common factor, the loadings of which are as follows:

a	*b*	*c*	*d*	*e*	*f*
.69	.69	.87	.87	.69	.69

No second common factor is indicated.

It should be noted that step 7 of the technique presents a serious obstacle in the procedure. Step "7" is as follows:

> Now make a list of all the tests that correlate positively with the test *b* and let this sub-group be designated *t*.[1]
>
> (Test "*b*" as used in this quotation is the second "pivot" test.)

When all the correlations of any matrix are positive the sub-group "*t*" will invariably include all of the tests in the matrix, and this, in turn, results in the conclusion that no second factor need ever be postulated in order to account for any matrix in which all the correlations are positive. (The particular weakness in the technique (step 7) which leads to this conclusion does not occur in Thurstone's more recently published methods of analysis.)

Furthermore, the first factor loadings, thus obtained, appear to be at strong variance with the facts as indicated by the correlation matrix. For instance, the first factor loading .69 in the hypothetical matrix, computed for each of tests *a*, *b*, *e* and *f*, conveys the idea that each of these tests contains a goodly portion of one and the same common factor, known as the First Factor. Yet a careful study of the correlation matrix itself reveals the fact that only a small amount of the loading .69 is due to a single common factor (tests *a* and *b* have the low correlation of .1 with tests *e* and *f*), and that most of it is due to the strong group factors underlying each pair of tests, *ab* and *ef* ($r_{ab} = .9$ and $r_{ef} = .9$). It thus appears that so-called First Factor Loadings are not necessarily representative of the loadings of any one common factor, but that they are more nearly representative of the relative average loadings of various factors.

It is also worthy of note that the results obtained do not nearly satisfy Thurstone's factor-equation, namely:

$$r_{ab} = a_1b_1 + a_2b_2 + a_3b_3 \cdots a_nb_n.$$

For instance, in the hypothetical case, when we substitute the appropriate values in this equation, we arrive at equations such as the following:

[1] *Psychological Review*, Sept., 1931, p. 423.

$$r_{ab} = .9 = .69 \times .69 + 0 = .48$$
$$r_{ae} = .1 = .69 \times .69 + 0 = .48$$

Let us now consider the results of analysis of the hypothetical matrix by Thurstone's "Simplified Multiple Factor Method."[1] The results obtained by this more recent method are more acceptable as far as the number of factors is concerned, but the factor loadings again appear to give a distorted picture of the amounts of each factor operating in the various tests. This method indicates that three common factors must be postulated in order to account for the correlations of this matrix. The loadings of each of these factors are as follows:

FACTORS[2]

Test	I	II	III	IV
a	+.678	+.664	−.004	.000
b	+.678	+.664	−.004	.000
c	+.858	−.113	−.389	.000
d	+.858	−.113	−.389	.000
e	+.678	−.541	+.377	.000
f	+.678	−.541	+.377	.000

These loadings satisfy the factor-equation.

$$r_{ab} = .9 = .678 \times .678 + .664 \times .664 + (-.004)(-.004) = .46 + .44 + .00 = .9$$
$$r_{ae} = .1 = .678 \times .678 + .664 \times (-.541) + (-.004 \times .337) = .46 - .36 - .00 = .1$$

However, an examination of the table of factor loadings reveals the existence of apparently serious limitations in the method of determining these loadings. The following points seem worthy of consideration.

1. When, as in this case, there are no negative correlations in the matrix, it seems difficult to interpret the meaning of negative loadings

[1] Thurstone, L. L.: *A Simplified Multiple Factor Method.* The University of Chicago Bookstore, Chicago, 1933, pp. 25.

[2] Step 18 of the simplified method requires that in the table of "First Factor Residuals," "the column with the highest sum disregarding sign" be selected. In this hypothetical case the columns *a*, *b*, *e*, and *f* are all equal and higher than the others. The loadings recorded above were obtained by assuming that the sums for columns *a* and *b* were greater than those for *e* and *f*. If *e* or *f* had been taken to be the greatest, the loadings now recorded for *a* and *b* should be exchanged with those for *e* and *f*.

in a way that would likely be generally acceptable; yet four of the six loadings, for each of the second and third factors, are negative.

2. It would appear that the loadings for tests *a* and *b* should be numerically the same as those for tests *e* and *f* (the second factor loadings for *a* and *b* should be the same as the third factor loadings for *e* and *f* and vice versa); yet the following marked differences occur:

	Factors	
Tests	II	III
a and *b*	+.664	−.004
e and *f*	−.541	+.337

3. It might be expected that tests *c* and *d* would show equal amounts of the second and third factors, but the loadings obtained are −.113 for the second factor and −.389 for the third factor.

4. The most serious of the difficulties encountered pertains to the first factor. The first factor loadings are practically the same as those obtained by the "1931" method. The same comments regarding them, therefore, apply.

Analysis by Thurstone's much more lengthy "Least Square" method, as published in 1932 in the "The Theory of Multiple Factors,"[1] results in almost the same first factor loadings as those obtained by the two methods just discussed. This long method, therefore, presents the same difficulty as the other methods in regard to the first factor.

In the opinion of the writer, it is doubtful if any method which has as a basic step in its procedure, the summation of the columns of correlations, will prove satisfactory.

As a further illustration of the points already noted we will consider one other case—one which shows even more clearly than the first that the practice of finding the sum of the correlations for each column is unlikely to prove fruitful.

In this matrix it is evident that no one factor is common to all the tests. It is obvious that three mutually independent group factors must be postulated in order to account for these correlations. Each test has the highest possible loading of one group factor and no

[1] Thurstone, L. L.: *The Theory of Multiple Factors.* Ann Arbor, Michigan, Edwards Brothers Inc., 1932, pp. 65.

loading of any other factor.[1] But the factor methods we have been discussing lead to quite different results. The "1931" and the "1933" methods each give a first factor loading of .58 for every variable, and the "1932" method gives a comparable, though slightly higher, first factor loading, .67. That is to say, that in this test case we find that whereas the first factor loading apparently should be 1.0 for each of two variables and .0 for each of the other four variables, the Thurstone methods of analysis give first factor loadings which are .58 or .67 for all variables.

HYPOTHETICAL CASE NO. 2

	a	*b*	*c*	*d*	*e*	*f*
a	...	1.0	0.0	0.0	0.0	0.0
b	1.0	...	0.0	0.0	0.0	0.0
c	0.0	0.0	...	1.0	0.0	0.0
d	0.0	0.0	1.0	...	0.0	0.0
e	0.0	0.0	0.0	0.0	...	1.0
f	0.0	0.0	0.0	0.0	1.0	...

In this particular test case, in which the sums of all the columns are equal, it is obviously correct to find that the loadings are all equal, but these are the loadings of three factors instead of one. It, therefore, appears that the techniques employed fail to differentiate the various factors present.

Some of the difficulties that have been mentioned, appear to be very important and should be overcome before interpretation of these techniques is clear.

---

[1] These group factors are not necessarily of a unitary character. Each might, conceivably, be composed of many factors. But, however that may be, the factor (or factors) underlying any one of the three pairs of variables, *ab*, *cd*, or *ef*, has nothing in common with the factor (or factors) underlying any other one of these pairs.

SOME ATTITUDES OF STUDENTS IN THE R.O.T.C.

HERBERT W. ROGERS

Lafayette College

Many of our colleges are challenging directly or indirectly the pertinence of the Reserve Officers Training Corps as a contributor to higher education. Recently a university has dropped the R.O.T.C., and several, including two state universities, have made the Basic R.O.T.C. course elective rather than compulsory. Student petitions to drop the R.O.T.C. or to have Compulsory R.O.T.C. made elective are becoming of more frequent occurrence. In institutions where the R.O.T.C. has been compulsory, individual petitioners for exemption have been numerous and in several institutions students have been expelled for refusing to drill. Several cases have been carried to court, and two of the Supreme Court of the United States.

Experimental studies of attitudes toward war and peace and of the R.O.T.C. have been rare. In 1926 Porter[6] compiled "Student Opinion of War," a test consisting of one hundred fifty questions. There were thirty-seven questions on the alleged causes of war, twenty-four questions on the results of war, forty-five questions on reducing the probability of war, thirteen questions on the right or wrong of war, eighteen questions on partiotism, and thirteen questions on American action in the immediate future. This test was given to one thousand students in eighteen institutions. Average scores for these institutions were obtained. Ninety-seven were retested a year later in which the average scores showed little change. Nineteen men undergoing military training during this time showed a very small shift toward pacifism. Twenty-seven men with such training, but not during the time between tests, showed a very small shift toward militarism. The correlation of the amount of military training and test scores of five hundred sixty-two men was $.3 \pm .03$. Average scores are presented for various religious denominations, for occupation of father, etc., but without populations and measures of variability. Group profiles were presented of Reserve Officers, American Legion officials, pacifists imprisoned during the war, etc. Porter concluded that questions of war and military training rise or lie dormant according to religious interest, social and economic liberalism, policy of R.O.T.C. staff, repressive policy of the school administration, pacifist literature,

speakers and the absence of other engrossing interests. The most potent cause of pacifism was religious conferences.

By the method of equal appearing intervals, Droba[2,3,4] has constructed an evenly graduated scale of militarism-pacifism. There are forty-four questions in the scale selected on the following topics; causes of war, results of war and peace, what is to be done now and in the event of war, and general judgments about war and peace. The reliability of the scale by the Spearman-Brown formula was .90. Droba presents evidence among two hundred fifty-five college students to show a consistent change toward pacifism in the successive undergraduate classes, and the graduate class. An r. of .15 (population two hundred twenty-seven) was found between militarism-pacifism and scholarship. Droba concludes from this *r.* that if any relationship exists there is a slight tendency for students of higher standing to be pacifists. Average scores on the scale were got for religious denomination but populations of the groups were too small to warrant the making of definite conclusions. Women are more pacifistic than men and vary less. Another group of one thousand students was used. Conclusions from this group were: College education is influential in modifying attitudes in the direction of pacifism. Women are more pacifistic and vary less. Neither intelligence (−.02) nor neurotic scores (Thurstone) (−.01) correlate with militarism-pacifism. Differences are found according to religious denomination. Catholics and Lutherans are the most militaristic, Jews and Christian Scientists the most pacifistic. There are other minor conclusions.

In 1932 Bishop[1] published a study of the value of military training. Questionnaires containing eight questions, with a letter from the then U. S. Commissioner of Education urging compliance, were sent to 16,068 college alumni who had been graduated from the R.O.T.C. course. This means that after taking two years in the Basic course (elective in twelve and compulsory in forty-two of the colleges covered) they elected to take the advanced course for their last two years and to go to a summer R.O.T.C. camp—all with the purpose of becoming R.O.T.C. officers. The group was highly selected. There was no control group. Only sixty-two per cent of the group returned the questionnaire. The questions were leading questions. Most of the questions were categorical. Conclusions are drawn which are not warranted by the returns. Opinion and attitude are construed to be accurate and conclusive facts. This study was financed by the "Mershon Foundation," whose operating subsidiary is the "Civilian

Military Education Fund." There was a second part of this pamphlet scheduled to appear but after the Education Commissioner's withdrawal from office this work was dropped due, in part, to numerous protests received. This study has been adversely criticized from various directions. Nothing may be gained by further criticising this study here.

In the last decade various statements have been made by military men about the purposes of the R.O.T.C. Nash[5] has listed, from the catalogs of state universities, thirty-eight of these stated objectives. They are too numerous and varied to discuss here. We are interested in this study, only in a few of the implicit purposes; in the attempt to build up a sympathy for the general idea of military training, of preparedness and the developing of a tolerance for the work and spirit of the regular army. In consequence of these purposes there is a probability that certain definite attitudes toward the Army, the R.O.T.C., war and nationalism, would be developed by the R.O.T.C. course. It is the purpose of this investigation to determine whether such are the facts.

In 1932 several students at Lafayette College who had an interest in Psychology tried out some questions among R.O.T.C. and Non-R.O.T.C. students in several colleges and universities. From this preliminary work our present list of questions was obtained. These questions were framed with care so that they would not encourage either a pro- or an anti-militaristic attitude or suggest any one answer to any question. In 1933 a group of students visited ten colleges and universities, and after obtaining permission of the R.O.T.C. commandant, they issued these forms to students and waited for them to be filled out. In many cases our investigators were granted permission to give the blanks to the R.O.T.C. members at their class period. The Control Group, *i.e.*, the men who had had no R.O.T.C., was got in an unselective manner by going from room to room in a dormitory, taking all men found in a fraternity house, or asking each man met on a campus to fill out the blank. In all our groups there were only two men who refused to fill in the answers. Hence we were able to overcome one chief cause of unreliability in the questionnaire method, that of obtaining only a part of a sample group.

The question sheet used is shown on pages 294 and 295.

Table I presents a tabulation of answers to the questionnaire. Our subjects were in two main groups, according to whether the R.O.T.C. was elective or compulsory in the institutions concerned. Each of

QUESTIONNAIRE

Course:

AB_____ BS_____ ENG_____ Other_____ Sign your name_____
(if you wish)

The aim of this survey is to learn something about human attitudes. Your cooperation in answering each question frankly and honestly will help us. Please answer *every* question.

On the bottom or back of this sheet—

Please say anything pro or con that you may wish.

You may expand any of your answers.

1. Was anyone in your family connected with military service?
   a. For duration of War 1917–1919_____. *c.* National Guard_____.
   b. As a permanent vocation _____. *d.* In no way _____.
2. Have you attended a military school? No_____ Yes_____ No. of Years_____.
3. How many years have you had R.O.T.C. in college (including this year)?
   0_____ 1_____ 2_____ 3_____ 4_____.
4. Has the R.O.T.C. altered your attitude toward the army?
   a. Unfavorably_____. *b.* Favorably_____. *c.* Not at all_____.
5. Do you think college students should be trained to assume military duties and responsibilities in times of peace? Yes_____. No_____.
6. Do you think the R.O.T.C. should be
   a. Dropped from the curriculum_____. *b.* Elective_____. *c.* Compulsory_____.
7. If required to wear a uniform throughout the week, would you (check one)
   a. Leave the course _____. *c.* Be annoyed _____.
   b. Register a complaint_____. *d.* Consider it a privilege_____.
   e. Find delight in wearing it_____.
8. The R.O.T.C. serves the following purposes (check all you think true).
   a. Develops character _____. *d.* Develops a feeling of self-confidence_____.
   b. Develops leadership_____. *e.* Trains for warfare _____.
   c. Develops patriotism_____. *f.* Serves no pertinent purpose _____.
9. Do you anticipate war (for the U. S.) within the next
   20 yrs._____ 40 yrs._____ 60 yrs._____ 80 yrs._____ 100 yrs._____ never_____.
10. After what period of time do you expect there will be no more wars in the civilized world?
    100 yrs._____ 200 yrs._____ 300 yrs._____ 400 yrs._____ 500 yrs._____ never_____.
11. In case of *offensive* war would you
    a. Volunteer_____. *b.* Wait for the draft_____. *c.* Evade the service_____.
    In case of *defensive* war would you
    a. Volunteer_____. *b.* Wait for the draft_____. *c.* Evade the service_____.
12. In time of war the best patriotism would make one
    a. Enlist, for it is the duty of every able-bodied citizen_____.
    b. Think for oneself and if war is thought wrong refuse to join_____.
    c. Think for oneself and if war is thought wrong enlist and when war is ended preach pacifism_____.

13. In case you did enter the army, during war, would you enter
    a. With a sense of duty _____. *c.* For the sake of adventure_____.
    b. For personal satisfaction_____. *d.* Other reason (state what)_____.
14. Do you think there is good to be derived from war?
    Yes_____. No_____. What good?_____
15. Do you think that diplomacy, humanistically enlightened, will avert
    a. In the near future: No wars_____ a few wars_____ many wars_____ most wars_____.
    b. In the remote future: No wars_____ a few wars_____ many wars_____ most wars_____.
16. Do you feel that Peace Conferences and Disarmament Conferences will avert
    a. In the near future: No wars_____ a few wars_____ many wars_____ most wars_____.
    b. In the remote future: No wars_____ a few wars_____ many wars_____ most wars_____.
17. Were any of the following elements responsible for your enrollment in the R.O.T.C.
    a. Interest in some special aspect.
    Horsemanship_____. Markmanship_____. Scabbard and Blade_____. Pay of cadets_____.
    b. Advice of
    Family_____. Teachers_____. Those giving loans for education_____.
    c. Any other reasons?_____
    Remarks:

these groups was sub-divided into a Basic Group, comprised of students in the first or second year, and an Advanced Group, those in the third or fourth year. A further sub-division in the Elective Group was a control section which had had no R.O.T.C.

In the Elective R.O.T.C. Group there were six institutions; two in the Boston area, four in the Pennsylvania, Maryland, Delaware area. In the Compulsory R.O.T.C. Group there were four institutions; one in the New York City area, and three in the Pennsylvania, Maryland, Delaware area. It was hoped that a more widespread Eastern area might be canvassed but this was not possible because of a lack of response on the part of some R.O.T.C. commandants. R.O.T.C. commandants in nineteen institutions were approached; and ten permitted us to do the work in their respective institutions.

At one institution we were fortunate in obtaining a Special Control Group. At this institution, where R.O.T.C. was compulsory, athletes were exempted from taking the course and these men formed the Special Control Group.

The population of the individual college groups ranged from twenty-two to eighty. This wide range was due to the variation in

TABLE I.—AVERAGE PERCENTAGE OF ELECTIVE AND COMPULSORY GROUPS

	Compulsory group					Elective group						Special control group	
			Average ratings						Average ratings				
	Advanced	Basic	Advanced	Basic	Weights	Advanced	Basic	Control	Advanced	Basic	Control	Control	Rating
Population	200	200	....	....	..	253	243	254	....	....	....	41	
Course.													
AB	23	39	....	....	..	23	32	39	....	....	....	22	
BS	19	28	....	....	..	47	43	30	....	....	....	41	
ENG	48	25	....	....	..	29	17	11	....	....	....	32	
Other	0	1	....	....	..	1	0	4	....	....	....	2	
No answer	10	7	....	....	..	0	9	16	....	....	....	5	
1. Family connection with military service.													
a. Duration War 1917–1918	35	32	....	....	..	34	42	27	....	....	....	41	
b. Permanent vocation	8	8	....	....	..	4	8	5	....	....	....	5	
c. National Guard	11	12	....	....	..	8	13	11	....	....	....	17	
d. No way	49	57	....	....	..	59	49	61	....	....	....	41	
No answer	4	1	....	....	..	0	1	3	....	....	....	7	
2. Attendance at Military School.													
No	92	86	....	....	..	92	91	95	....	....	....	93	
Yes	6	10	....	....	..	8	7	3	....	....	....	5	
No answer	1	5	....	....	..	0	3	2	....	....	....	2	
3. Years of ROTC in college.													
0	0	0	....	....	..	0	0	100	....	....	....	100	
1	0	27	....	....	..	0	47	0	....	....	....	0	
2	0	73	....	....	..	0	53	0	....	....	....	0	
3	44	0	....	....	..	56	0	0				0	
4	57	0	....	....	..	45	0	0	....	....	....	0	
4. ROTC altered attitude to army.													
a. Unfavorably	3	12	....	....	..	5	11	11	....	....	....	15	
b. Favorably	61	37	....	....	..	70	51	11	....	....	....	5	
c. Not at all	37	51	....	....	..	24	38	50	....	....	....	54	
No answer	0	1	....	....	..	1	0	28	....	....	....	27	
5. Should students be trained to assume military duties and responsibilities in times of peace.													
Yes	92	70	....	....	..	91	80	34	....	....	....	34	
No	7	30	....	....	..	8	18	65	....	....	....	63	
No answer	2	1	....	....	..	1	3	2	....	....	....	2	
6. Think ROTC should be													
a. Dropped from curriculum	0	9	....	....	3	1	8	23	....	....	....	22	
b. Elective	50	62	....	....	2	78	71	71	....	....	....	73	
c. Compulsory	50	27	....	....	1	21	22	5	....	....	....	5	
No answer	0	2	....	....	..	0	0	1	....	....	....	0	
Average rating	...	...	1.50	1.83	..	...	...	...	1.79	1.86	2.19	...	2.17
7. Wear uniform throughout week.													
a. Leave course	1	15	....	....	..	3	20	27	....	....	....	22	
b. Register complaints	18	27	....	....	..	22	16	16	....	....	....	37	
c. Be annoyed	35	41	....	....	..	23	32	41	....	....	....	32	
d. Consider it privilege	38	6	....	....	..	34	22	6	....	....	....	0	

TABLE I.—*Continued*

	Compulsory group					Elective group						Special control group	
			Average ratings						Average ratings				
	Advanced	Basic	Advanced	Basic	Weights	Advanced	Basic	Control	Advanced	Basic	Control	Control	Rating
e. Delight in wearing it	3	1	....	....	..	9	2	4	....	....	....	5	
Indifferent	3	4	....	....	..	6	3	1	....	....	....	0	
No answer	4	2	....	....	..	4	4	5	....	....	....	5	
8. Think ROTC serves following purposes.[1]													
a. Develops character	68	34	....	....	..	70	51	27	....	....	....	7	
b. Develops leadership	95	65	....	....	..	94	77	38	....	....	....	24	
c. Develops patriotism	53	49	....	....	..	64	55	39	....	....	....	24	
d. Develops a feeling of self-confidence	93	35	....	....	..	80	57	22	....	....	....	15	
e. Trains for warfare	70	64	....	....	..	71	77	63	....	....	....	54	
f. Serves no pertinent purpose	0	19	....	....	..	1	13	27	....	....	....	41	
No answer	0	1	....	....	..	0	1	1	....	....	....	2	
9. Anticipate war (for U. S.) within													
20 years	68	75	....	....	1	60	63	51	....	....	....	56	
40 years	14	11	....	....	2	19	17	22	....	....	....	12	
60 years	1	2	....	....	3	2	5	6	....	....	....	15	
80 years	0	1	....	....	4	1	1	4	....	....	....	0	
100 years	3	3	....	....	5	4	3	6	....	....	....	0	
Never	2	5	....	....	6	4	4	7	....	....	....	15	
No answer	14	3	....	....	..	11	7	5	....	....	....	2	
Average rating	...	...	1.34	1.53	..	...	...	...	1.63	1.70	2.10	...	2.19
10. No more wars in civilized world after													
100 years	4	7	....	....	6	8	8	17	....	....	....	17	
200 years	1	2	....	....	5	5	5	9	....	....	....	7	
300 years	2	3	....	....	4	1	2	4	....	....	....	2	
400 years	0	0	....	....	3	0	1	1	....	....	....	0	
500 years	6	6	....	....	2	1	9	6	....	....	....	5	
Never	79	80	....	....	1	75	67	56	....	....	....	63	
No answer	10	4	....	....	..	10	9	4	....	....	....	5	
Average rating	...	...	1.39	1.57	..	...	...	...	1.78	1.81	2.40	...	2.29
11. In case of *offensive* war would													
a. Volunteer	35	25	....	....	1	44	25	18	....	....	....	10	
b. Wait for draft	54	49	....	....	2	45	51	41	....	....	....	41	
c. Evade service	6	25	....	....	3	8	21	40	....	....	....	49	
No answer	6	2	1.69	1.99	..	3	4	2	1.64	1.97	2.23	0	2.39
In case of *defensive* war would													
a. Volunteer	81	69	....	....	1	87	69	49	....	....	....	37	
b. Wait for draft	16	23	....	....	2	12	22	30	....	....	....	32	
c. Evade service	0	9	....	....	3	0	6	18	....	....	....	29	
No answer	4	1	....	....	..	2	2	3	....	....	....	2	
Average rating	...	...	1.16	1.42	...	...	...	...	1.12	1.34	1.68	...	1.92

[1] Percentages total more than one hundred since question called for more than one item to be marked.

TABLE I.—*Continued*

	Compulsory group					Elective group						Special control group	
			Average ratings						Average ratings				
	Advanced	Basic	Advanced	Basic	Weights	Advanced	Basic	Control	Advanced	Basic	Control	Control	Rating
12. In time of war best patriotism would make one													
a. Enlist—duty of every able-bodied citizen	56	43	....	....	1	63	46	31	....	....	....	32	
b. Think—if war thought wrong refuse to join	20	32	....	....	3	12	28	43	....	....	....	57	
c. Think—if war thought wrong enlist, when war is ended preach pacifism	15	23	....	....	2	18	19	21	....	....	....	17	
No answer	10	2	....	....	..	8	7	4	....	....	....	0	
Average rating	...	...	1.60	1.89	..	...	...	...	1.47	1.81	2.12	...	2.37
13. Entering army during war for													
a. Sense of duty	85	67	....	....	..	86	78	54	....	....	....	37	
b. Personal satisfaction	7	6	....	....	..	11	9	5	....	....	....	10	
c. Sake of adventure	19	26	....	....	..	13	15	20	....	....	....	32	
d. Other reason	6	3	....	....	..	5	5	9	....	....	....	5	
(*a*) Social pressure	0	0	....	....	..	0	2	1	....	....	....	0	
(*b*) Compulsion	1	6	....	....	..	1	4	17	....	....	....	24	
(*c*) Safety	1	0	....	....	..	1	1	4	....	....	....	0	
No answer	4	4	....	....	..	1	3	1	....	....	....	0	
14. Good derived from war.													
Yes	26	33	....	....	..	32	28	23	....	....	....	15	
No	67	67	....	....	..	65	68	73	....	....	....	85	
No answer	7	1	....	....	..	3	4	5	....	....	....	0	
What good													
a. Protection	3	2	....	....	..	3	2	3	....	....	....	2	
b. Speeds scientific work	1	2	....	....	..	3	1	0	....	....	....	2	
c. Checks over-population	4	1	....	....	..	2	4	3	....	....	....	0	
d. Makes prosperity	0	0	....	....	..	0	0	1	....	....	....	0	
15. Diplomacy will avert													
a. In near future													
No wars	11	13	....	....	1	11	11	16	....	....	....	10	
Few wars	49	52	....	....	2	49	51	43	....	....	....	32	
Many wars	17	11	....	....	3	14	13	15	....	....	....	24	
Most wars	18	23	....	....	4	21	20	19	....	....	....	29	
No answer	6	2	....	....	..	5	5	8	....	....	....	5	
Average rating	...	...	2.45	2.46	..	...	...	...	2.47	2.44	2.39	...	2.76
b. In remote future													
No wars	6	6	....	....	1	9	13	16	....	....	....	12	
Few wars	28	34	....	....	2	32	23	21	....	....	....	15	
Many wars	18	18	....	....	3	18	21	19	....	....	....	17	
Most wars	39	38	....	....	4	34	33	34	....	....	....	44	
No answer	9	4	....	....	..	8	9	10	....	....	....	12	
Average rating	...	...	2.96	2.93	..	...	...	...	2.83	2.81	2.79	...	3.06

TABLE I.—*Continued*

	Compulsory group					Elective group						Special control group	
			Average ratings						Average ratings				
	Advanced	Basic	Advanced	Basic	Weights	Advanced	Basic	Control	Advanced	Basic	Control	Control	Rating
16. Peace and Disarmament Conferences will avert													
a. In near future													
No wars	32	28	....	....	1	28	29	25	....	....	....	27	
Few wars	51	50	....	....	2	51	44	46	....	....	....	49	
Many wars	8	12	....	....	3	8	10	12	....	....	....	5	
Most wars	4	10	....	....	4	10	11	10	....	....	....	17	
No answer	5	1	....	....	..	3	6	7	....	....	....	2	
Average rating	...	...	1.83	2.05	..	...	...	...	2.00	2.03	2.07	...	2.15
b. In remote future													
No wars	23	15	....	....	1	20	25	22	....	....	....	15	
Few wars	39	45	....	....	2	41	31	26	....	....	....	29	
Many wars	14	17	....	....	3	11	17	17	....	....	....	22	
Most wars	15	22	....	....	4	22	21	24	....	....	....	29	
No answer	10	2	....	....	..	6	6	11	....	....	....	5	
Average rating	...	...	2.22	2.47	..	...	...	...	2.38	2.37	2.46	...	2.68
17. Were any of the following elements responsible for your enrollment in the ROTC?													
a. Interest in some special aspect.													
Horsemanship	5	2	....	....	..	7	12						
Marksmanship	16	11	....	....	..	9	18						
Scabbard and Blade (The Military Fraternity)	4	3	....	....	..	9	6						
b. Advice of													
Family	11	13	....	....	..	20	17						
Teachers	7	7	....	....	..	9	6						
c. Other reasons													
Interest in military training	33	6	....	....	..	35	21						
Compulsion	1	39	....	....	..	1	2						
Defense	4	0	....	....	..	2	1						
To gain self-confidence	10	0	....	....	..	5	2						
Fellowship	6	0	....	....	..	0	2						
Preference to gym	0	0	....	....	..	3	8						
Commission	5	1	....	....	..	4	4						
Leadership	13	2	....	....	..	11	2						
Chances less of being killed if one were an officer rather than a private	7	2	....	....	..	3	1						
Discipline	2	0	....	....	..	1	3						
Character training	2	0	....	....	..	8	1						
Academic credit	1	1	....	....	..	3	4						
Other answers	14	3	....	....	..	18	15						
No answer	3	28	....	....	..	3	9						

the number of men taking the Advanced Course. Our method, in the colleges where the group was small, was to interview all or nearly all (never less than eighty-eight per cent) of the men in the Advanced Course and then to obtain an equal number in the Basic Course and, in institutions where R.O.T.C. was elective, to get an equal number in the Control Group. No Control Group was possible where R.O.T.C. was compulsory except in the one institution where athletes were exempted from the course. Since the populations of the groups varied considerably and in order that the larger groups might not warp the smaller groups, the per cent of answers is taken for each institution and these per cents are averaged.

The answer to some questions permitted a quantitative rating and the column headed "weights" gives the weightings assigned to answers. In each case this is a simple progression and is in the order militarism-pacifism. The weight 1 being the attitude most favorable to militarism and the largest number the attitude most favorable to pacifism. Significant results seems to be:

Q. 2. Have you attended a military school?

In the colleges where R.O.T.C. is elective, 2½ students who have attended a military school elect R.O.T.C. for each student who does not elect it.

Q. 4. Has the R.O.T.C. altered your attitude toward the army—unfavorably, favorably, not at all?

Almost two-thirds of the men in the R.O.T.C. think that their attitude toward the army has altered; becoming more favorable in fifty-five per cent and less favorable in eight per cent. The Control Group is unaffected; the number of those who have acquired a more favorable attitude toward the army is about equaled by those whose attitude has become less favorable.

Q. 5. Do you think college students should be trained to assume military duties and responsibilities in times of peace?

Three times as many Basic students think students should be trained for war in times of peace, as compared to the number of Basic students who think they should not be so trained. There is a reliable difference between the Compulsory and Elective Groups—the above tendency is less marked among the Basic students, of the former group. Is this due to the fact that in this group those who dislike R.O.T.C. must take it? The tendency is pronounced in the Advanced

Course. For every twelve Advanced Course students who think that students should be trained for war in times of peace there is but one Advanced Course student who thinks not. The opposite tendency is observed in the Control Group; for each student who thinks well of this training for war in times of peace, there are two who think he should not be so trained.

Q. 6. Do you think the R.O.T.C. should be—dropped from curriculum, elective, compulsory?

For every R.O.T.C. student in the Advanced Course who thinks that R.O.T.C. should be compulsory there are two who think it should be elective. For every Basic R.O.T.C. student who thinks Compulsory R.O.T.C. desirable there are three who think Elective R.O.T.C. is preferable. In both Control Groups, for every man who thinks the course should be compulsory there are fourteen who think not. Eight or nine per cent of the Basic students think that such work has no place in college. In the Control Group this per cent rises to twenty-three. The weighting of answers is used for the first time in Question 6. This treatment of the results gives a relative figure for each of the groups, Advanced, Basic and Control. In the Elective Group there is a difference of .07 points between the average rating of the advanced and basic sections (1.79–1.86). The standard deviations of the distributions are .42 and .53; the standard deviations of the averages are .027 and .034 and the standard deviation of the difference is .043. The chances that this difference, in the two averages, is not due to chance are nine hundred forty-seven in one thousand. Respective measures for the Basic and Control sections are difference in average .33, σ distribution .53 and .50, σ average .034 and .031, σ difference .046, and there one thousand chances in one thousand that the difference in the two averages is not due to chance. With these figures as reference points it may be noted by observation, that the difference in the averages of the Advanced and the Control sections in the Elective Group (.40) is reliable and that the difference in the averages of the Advanced and Basic sections in the Compulsory Group (.33) is reliable. It is a waste of good time to do the arithmetic in most of the averages which follow. Sigma differences will be obtained only in cases where, from observation, it may be anticipated that the difference in the average is not highly reliable. Table II gives these measures.

TABLE II.—MEASURES OF RELIABILITY OF DIFFERENCES IN AVERAGES

	Averages	Difference in average	$\sigma_{diff.}$	Chances in one thousand
Q. 6. Elective Group, Advanced and Basic	1.79 1.86	.07	.043	947
Q. 6. Elective Basic—Control	1.86 2.19	.33	.046	1000
Q. 9. Elective Advanced—Basic	1.63 1.70	.07	.136	702
Q. 10. Elective Advanced—Basic	1.78 1.81	.03	.172	568
Q. 15*a*. Compulsory	2.45 2.46	.01	.095	544
Q. 15*a*. Elective Advanced—Basic	2.47 2.44	.03	.098	621
Q. 15*a*. Elective Basic—Control	2.44 2.39	.05	.100	688
Q. 15*a*. Elective Advanced—Control	2.47 2.39	.08	.101	785
Q. 15*b*. Compulsory	2.96 2.93	.03	.103	614
Q. 15*b*. Elective Advanced—Basic	2.83 2.81	.02	.111	571
Q. 15*b*. Elective Basic—Control	2.81 2.79	.02	.117	568
Q. 15*b*. Elective Advanced—Control	2.83 2.79	.04	.113	637
Q. 16*a*. Elective Advanced—Basic	2.00 2.03	.03	.094	626
Q. 16*a*. Elective Basic—Control	2.03 2.07	.04	.096	663
Q. 16*a*. Elective Advanced—Control	2.00 2.07	.07	.092	776
Q. 16*b*. Elective Advanced—Basic	2.38 2.37	.01	.112	536
Q. 16*b*. Elective Basic—Control	2.37 2.46	.09	.116	782
Q. 16*b*. Elective Advanced—Control	2.38 2.46	.08	.115	758

Q. 7. If required to wear a uniform throughout the week, would you—leave the course, register a complaint, be annoyed, consider it a privilege, find delight in wearing it?

The uniform plays its part. There are marked degrees of difference between groups. The Special Control Group is more vehement in its objection than other Control Groups.

Q. 8. The R.O.T.C. serves the following purposes—develops character, develops leadership, develops patriotism, develops a feeling of self-confidence, trains for warfare, serves no pertinent purpose (check all you think true).

Here again there are marked differences between the returns from the respective groups. The more R.O.T.C. training a man has the more likely is he to think that there is a variety of personality development resulting from taking R.O.T.C. The development of character, of leadership and of self-confidence by R.O.T.C. training are all accepted in direct proportion to the amount of R.O.T.C. a student has had. More than half of the R.O.T.C. students believe it develops patriotism. More than two-thirds believe it trains for warfare. This is the only element of Question 8 which is answered affirmatively by

more than half of the Controls. Again the Special Control Group findings are more positive than the other Control Group.

Q. 9. Do you anticipate war (for the U. S.) within the next—twenty, forty, sixty, eighty, one hundred years; never?

Q. 10. After what period of time do you expect there will be no more wars in the civilized world—one hundred, two hundred, three hundred, four hundred, five hundred years; never?

The anticipation of war varies significantly with the Compulsory Group. In the Elective Group there is a significant difference only between the Control Group and the other groups.

Q. 11. In case of Offensive war would you—volunteer, wait for the draft, evade the service?

In case of Defensive war would you—volunteer, wait for the draft, evade the service?

Estimates of one's behavior about enlistment in the event of war show marked differences between all groups. May we point out again that we are not here interested in what these men would actually do in the event of another war but we are interested simply in their present attitudes?

Q. 12. In time of war the best patriotism would make one—enlist, for it is the duty of every able-bodied citizen; think for oneself and if war is thought wrong refuse to join; think for oneself and if war is thought wrong enlist and when war is ended preach pacifism.

Q. 13. In case you did enter the army, during war, would you enter—with a sense of duty, for personal satisfaction, for the sake of adventure, other reason (state what)?

Marked differences are noted between groups in the attitude toward enlistment in the event of war and motives for enlistment.

Q. 14. Do you think there is good to be derived from war? What good?

No significant differences are found except in the case of the Special Control Group.

Q. 15. Do you think that diplomacy, humanistically enlightened, will avert—(*a*) in the near future: No wars, a few wars, many wars, most wars; (*b*) in the remote future: No wars, a few wars, many wars, most wars.

Q. 16. Do you feel that Peace Conferences and Disarmament Conferences will avert—(*a*) in the near future: No wars, a few wars, many wars, most wars; (*b*) in the remote future: No wars, a few wars, many wars, most wars.

No differentiating result except for the Special Control Group in relation to all other groups.

Q. 17. Were any of the following elements responsible for your enrollment in the R.O.T.C.—(*a*) Interest in some special aspect: Horsemanship, marksmanship, Scabbard and Blade, pay of cadets; (*b*) Advice of: Family, teachers; (*c*) other?

There is no control group for this question since the question is pertinent only to those who have had R.O.T.C. There are however interesting differences between the Advanced and Basic Groups. Some of the extra features of the R.O.T.C., such as pay and marksmanship loom large.

CONCLUSIONS

To all of the questions asked, except Questions 14, 15 and 16, answers were given which, on the average, indicate differences of attitude between the Advanced, Basic and Control Groups. In a few cases there were differences between only two of these three groups, but in most of the questions there was a reliable line for line difference in all the groups. The lack or amount of R.O.T.C. training which a man has varies directly with his attitude toward certain relevant questions. Knowing that a group of college men has not had R.O.T.C. training, or knowing the amount of training that it has had, we may prognosticate its attitudes, such as; sympathy toward the army, approval of military training in times of peace, opinion as to whether R.O.T.C. should be compulsory or elective or abolished, response to the glamour of the uniform, acceptance of the R.O.T.C. as a developer of certain traits of character, the anticipation of war, the hope of a pacific world, ideas of desirable behavior in the event of another war, belief in "patriotism" in the sense of unquestioning obedience to mandates to fight for one's country in the event of another war and motives for so doing.

A further question may be asked, whether the R.O.T.C. training has produced these attitudes or whether the R.O.T.C., where that course was elective, has merely selected the men who had previously acquired these attitudes, and further selected among their basic students the men who had these attitudes in a pronounced form, and in compulsory institutions had selected the most zealous for the Advanced Course. We have no way of knowing from these results the answer to this question. At least, it is evident from our returns that if the men started the course with these varying attitudes the R.O.T.C. did little to change them.

The returns from some questions lead one to expect that the R.O.T.C. is building up these attitudes. In one institution eleven,

or fourteen per cent, of the Advanced Course students state that war serves a good in that it accelerates scientific progress. In the other nine institutions a total of but ten students gave this answer. In one institution one hundred per cent of both Advanced and Basic Groups state that they think that the training develops leadership. In Question 4, fifty per cent of the Control Group state that they feel they have been uninfluenced in their attitude toward the army. Of those taking R.O.T.C. almost two-thirds think that they have been influenced and of these the great majority have been influenced favorably. If this occurs in this instance it is highly probable that there are other directions in which attitudes are developed. The attitude toward the status of the R.O.T.C.—whether it should be compulsory or elective or abolished—may be largely due to the work of the R.O.T.C. since most students before coming to college have probably had little interest in such a question. The widespread feeling (Question 8) among R.O.T.C. students that their course develops self-confidence raises some interesting questions. Does a feeling of inferiority or a feeling of personal insecurity lead one to elect R.O.T.C.? Having had the Basic Course, does such a feeling motivate one to take more R.O.T.C.? Question 12, referring to effects of patriotism upon enlistment, shows that R.O.T.C. men believe, that in the event of another war, "patriotic" motives would be a strong factor in motivating their enlistment. Since there are reliable differences in all groups there is an indication that either the R.O.T.C. is instilling obedience to an unthinking kind of patriotism, or the R.O.T.C. selects and holds men already having this idea of patriotism. Answers to Question 14, "is there a good derived from war?" show two-thirds of all men seem convinced of the futility of war. Very few of those stating that war resulted in some good were willing to specify the good—only one man in four so doing. From a study of individual institutions it is found that where a specific good was stated there was a piling up of the same answer in a single institution, strongly suggesting the result of indoctrination.

BIBLIOGRAPHY

1. Bishop, Ralph C.: *A Study of the Educational Value of Military Instruction in Universities and Colleges.* Office of Education, U. S. Department of the Interior, Washington, D. C., 1932.
2. Droba, D. D.: *A Scale for Measuring Attitude Toward War.* University of Chicago Press, 1930.

3. Droba, D. D.: "The Effect of Various Factors on Militarism-Pacifism." *J. Abnormal and Social Psychology*, 1931, Vol. XXVI, pp. 131–153.
4. Dorba, D. D.: "A Scale of Militarism-Pacifism." *J. Educational Psychology*, Vol. XXII, 1931, pp. 96–111.
5. Nash, Willard, H.: *The Stated Aims and Purposes of the Departments of Military Science and Tactics, and Physical Education in the Land Grant Colleges of the United States.* Contributions to Education, Teachers College, Columbia University, 1934.
6. Porter, Eliot: *Student Opinion on War: An Investigation.* Unpublished thesis in the University of Chicago Library, 1926.

A CORRELATIONAL METHOD APPLICABLE TO THE STUDY OF THE TIME FACTOR IN INTELLIGENCE TESTS

M. E. ODOROFF

University of Minnesota

Since May's[1] study on the "Effect of Doubling the Time Limits in the Alpha and Beta Examinations," a number of investigators have studied the effect of increasing the time limits on intelligence test performance by correlating scores made under standard time limits with scores made under extended time. A few of the coefficients reported were very high[2] and were interpreted to mean that an individual's relative position in the group was not greatly altered by the increase in time. Other coefficients[3] were lower and were interpreted as meaning that standard time limits handicapped some individuals by preventing them from attaining their relative level of ability.

A survey of the literature revealed that only one of the investigators[4] recognized a difficulty inherent in the experimental technique. The difficulty arises from the fact that, if a testing technique like May's[5] is used, the standard time score is either part or the whole of the extended time score. If, however, the experimental technique used is a test-retest procedure similar to Freeman's,[6] the same difficulty, though somewhat reduced, still remains, since up to the point where the time limit is being extended, the correlation between the two

---

[1] May, M.: "Effect of Doubling the Time Limits in the Alpha and Beta Examinations." *Memoirs Nat. Acad. Sci.*, Vol. XV, 1921, pp. 415–420.

[2] May, M.: *Op. cit.*, Ruch, G. M., and Koerth, W.: " 'Power' *vs.* 'Speed' in Army Alpha." *J. of Educ. Psychol.*, Vol. XIV, 1923, pp. 193–208. Ruch, G. M.: "The Speed Factor in Mental Measurements." *J. of Educ. Research*, Vol. IX, 1924, pp. 39–45.

[3] Freeman, F. S.: "Power and Speed: Their Influence upon Intelligence Test Scores." *J. of Appl. Psychol.*, Vol. XII, 1928, pp. 631–635. Freeman, F. S.: "The Factors of Speed and Power in Tests of Intelligence." *J. of Exper. Psychol.*, Vol. XIV, 1931, pp. 83–90. Freeman, F. S.: "The Factor of Speed." *J. of Gen'l Psychol.*, Vol. VI, 1932, pp. 462–468. Walters, F. C.: "A Statistical Study of Certain Aspects of the Time Factor in Intelligence." *Teachers Coll. Contrib. to Educ.*, No. 248, 1927, pp. 82. Tinker, M. A.: "Influence of the Speed Attitude on Test Performance." *J. of Gen'l Psychol.*, Vol. X, 1934, pp. 465–469.

[4] Ruch and Koerth: *Op. cit.*

[5] May: *Op. cit.*

[6] Freeman: *Op. cit.*

variables is merely a reliability coefficient. As a result, the correlation between the standard and extended time scores will be overestimated, and the apparent relationship between the two variables is spurious. This paper suggests a correlational method borrowed from the field of biology and derived by Harris[1] that avoids the difficulty.

The method consists in determining whether the dependent variable becomes relatively larger with increase or decrease in the independent variable, and is arrived at by correlating the independent variable with the deviation of the dependent variable from its probable value. A positive value of the coefficient would indicate that those scoring high under standard time increase their score relatively more than those scoring low. A negative value of the coefficient would indicate that those scoring low under standard time increase their score relatively more than those scoring high, and a zero coefficient would indicate a proportional increase in score for all of the subjects, since the ranks earned under standard time are retained under extended time.

The Otis Self-Administering Tests of Mental Ability, Higher Form *A* was administered to ninety-two subjects. Forty-seven were members of a class in Educational Psychology at the University of Minnesota. The other forty-five were members of a class in Elementary Psychology in an evening school class in the Adult Education division of the Minneapolis Public Schools. At the end of twenty minutes, colored pencils were distributed, and the subjects were told to continue until every item had been attempted at least once. They were also told that they might correct any item already attempted.

TABLE I.—MEANS, STANDARD DEVIATIONS, AND COEFFICIENTS OF VARIATION FOR STANDARD AND EXTENDED TIME SCORES FOR THE OTIS SELF-ADMINISTERING TESTS OF MENTAL ABILITY, HIGHER FORM *A*

	Mean	SD	CV
Standard time	48.3	9.1	18.8
Extended time	53.8	10.5	19.5

Means, standard deviations and relative variation for both distributions are given in Table I. Extension of the time limits yielded

[1] Harris, J. A.: "The Correlation of a Variable and the Deviation of a Dependent Variable from Its Probable Value." *Biometrika*, Vol. VI, 1909, pp. 438–443.

little increase in score. The relative variation for the two distributions is approximately equivalent.

The Pearson r between the standard and the extended time scores is $.905 \pm .013$. The r computed by the Harris[1] formula is $-.297 \pm .064$. The latter coefficient is interpreted as meaning that there is a slight tendency for the subjects scoring low under standard time limits to increase their scores relatively more than those scoring high. The standard time limits handicap a few individuals by preventing them from achieving their relative level of ability.

Since the Harris r seems to have some merit as a method of measuring the effect of the time factor in intelligence, the writer applied it to the data found in the literature. Table II presents the mean scores, standard deviations and comparative correlations for the standard and extended time[2] scores of several intelligence tests.

An examination of the data in Table II emphasizes the utility of the Harris formula. The Pearson r's for the same tests vary considerably, and as a result have been variously interpreted. On the other hand, the Harris r's lend themselves to interpretation more easily. For instance, the variation in the size of the coefficients for the Army Alpha is explained by noting the mean scores and the variabilities of the group concerned. The different levels of ability at which these groups work account for the difference in the size of the coefficients. For a group such as May[3] tested, the Army Alpha is difficult enough to be classified as a "power" test. The evidence bears out this conclusion. The Harris r for May's group is .136. It is interpreted to mean an individual retains under extended time the relative position earned under standard time.

An inspection of the results for Ruch and Koerth's data leads one to conclude that the Army Alpha is a "speed" test for the more intelligent college student. Ruch and Koerth[4] selected their subjects

---

[1] The Harris formula is:

$$r_{xx} = \frac{r_{xy} - \frac{v_x}{v_y}}{\sqrt{1 - r^2_{xy} + \left(r_{xy} - \frac{v_x}{v_y}\right)^2}}$$

r_{xy} is the Pearson product moment coefficient, and v_x and v_y are the coefficients of variation for the extended and standard time scores respectively.

[2] The x variable is the extended time score.

[3] May: *Op. cit.*

[4] Ruch and Koerth: *Op. cit.*

from the highest and lowest deciles on the basis of the combined scores of four intelligence tests previously administered. Since of their one hundred twenty-two subjects, fifty-two were in the highest decile, the above conclusion seems justified. The Harris r's for their

TABLE II.—MEAN SCORES, STANDARD DEVIATIONS AND COMPARATIVE COEFFICIENTS FOR THE STANDARD AND EXTENDED TIME SCORES OF SEVERAL INTELLIGENCE TESTS

Author and test	M_x	M_y	SD_x	SD_y	Pearson r_{xy}	Harris r_{xx}	N
May.							
Army alpha[1]	80.5	62.0	42.2	35.0	.965	.136	510
Ruch and Koerth.							
Army alpha	155.0	127.6	31.9	38.2	.945	.615	122
Army alpha[1]	149.6	127.6	34.4	38.2	.966	.604	122
Terman	97.9	82.1	32.1	30.7	.960	.284	150
Walters.							
N. I. T.	120.4	101.9	18.1	17.6	.883	.034	162
Otis advanced	103.7	87.8	23.0	19.8	.912	−.168	157
Cross out	60.1	56.6	9.6	9.7	.944	.024	155
Pintner non-language	408.2	384.7	47.5	51.0	.831	−.086	158
Freeman.[2]							
N. I. T.	.....	.....	.....	.....	.83	.05	56
Otis advanced	.....	.....	.....	.....	.58	−.29	51
Ohio State University psychological examination	223.7	134.1	29.5	34.2	.72	.28	117
Terman	.....	.....	.....	.....	.93	.35	42
Dearborn[1]	44.5	36.1	12.6	11.5	.88	−.02	100
Tinker.							
Army alpha[1] (Form 5)	165.1	128.1	13.1	17.2	.786	.302	110
Army alpha[1] (Form 9)	166.6	129.3	12.5	18.5	.719	.270	110
Army alpha (Form 5)	173.2	133.3	14.1	20.2	.785	.377	111
Army alpha (Form 9)	171.6	127.6	14.2	22.8	.746	.390	111

[1] The extended time limits for these tests are merely double the standard time.

[2] Mean scores and variabilities were not published for the N. I. T., Otis Advanced and the Terman tests.

group are .615 and .604 for standard time with extended and double time respectively. The trend is the same for Tinker's college group.

The coefficients for the remaining tests, with the possible exceptions of the Terman Group Test and the Ohio State University Psychological Examination, cluster around zero. They are interpreted

as meaning that an increase in time limits results in little change in the relative level of performance.

SUMMARY

The present study has pointed out that the Pearson product moment coefficient of correlation overestimates the relationship between the standard and the extended time scores. This spuriousness is due to the fact that the standard time score is either part or the whole of the extended time score. A correlational method previously published by Harris, that may be used in determining the direction of the relative increase in score due to the extension of the time limits, was suggested in place of the Pearson *r*. Application of the Harris *r* to the writer's own data and the data taken from the literature leads one to conclude that, within its limitations, it is a better technique for this and similar problems[1] than the Pearson *r*.

[1] Eurich, A. C.: "Retention of Knowledge Acquired in a Course in General Psychology." *J. of Appl. Psychol.*, Vol. XVIII, 1934, p. 217.

THE RELIABILITY OF INCREMENTS

FRANK K. SHUTTLEWORTH

Yale University

From a number of sources, scattered both as to time and field of interest, the suggestion has been advanced that our program of measuring and testing children places too much emphasis on status or achievement or ability at a given time. It has been urged that measurements and tests should be repeated at quarterly or semi-annual or annual intervals and that the emphasis should fall on growth, progress, gains in achievement, increase in ability, and similar dynamic concepts. From this point of view the important fact is not that Johnny's health is below par, the important consideration may be that his health has improved markedly over the past year. Similarly, if Mary's educational achievement is above average for her age, the important consideration may be that she has made no progress during the interval of a year and now has a percentile rank of sixty instead of ninety. If this point of view is to function, the question arises as to the reliability with which individual differences in gains or increases or increments can be measured. The purpose of this paper is to present a formula for determining this reliability.

An increment of physical growth, educational achievement, or ability may be defined as

$$d_1 = y_1 - x_1 = (y_\infty + e_{y_1}) - (x_\infty + e_{x_1})$$

in which d_1 is the difference or increment, x_1 is the initial measurement, y_1 is the second measurement obtained at an interval of months or years, x_∞ and y_∞ are true scores, and e_{x_1} and e_{y_1} are errors of measurement. Repetition of these measures or the application of comparable forms of the test gives a second measurement of the increment

$$d_2 = y_2 - x_2 = (y_\infty + e_{y_2}) - (x_\infty + e_{x_2})$$

The correlation $r_{d_1 d_2}$, the reliability with which individual differences in increments is measured, is desired. Multiplying the above equations, summing, and dividing by N gives

$$\frac{\Sigma d_1 d_2}{N} = \frac{\Sigma(y_\infty + e_{y_1} - x_\infty - e_{x_1})(y_\infty + e_{y_2} - x_\infty - e_{x_2})}{N}$$

On the assumption that the errors are random, this reduces to

$$\frac{\Sigma d_1 d_2}{N} = \frac{\Sigma y_\infty{}^2}{N} + \frac{\Sigma x_\infty{}^2}{N} - \frac{2\Sigma y_\infty x_\infty}{N}$$

and

$$r_{d_1 d_2}\sigma_{d_1}\sigma_{d_2} = \sigma_{y_\infty}{}^2 + \sigma_{x_\infty}{}^2 - 2r_{y_\infty x_\infty}\sigma_{y_\infty}\sigma_{x_\infty}$$

and

$$r_{d_1 d_2}\sigma_{d_1}\sigma_{d_2} = r_{y_1 y_2}\sigma_y{}^2 + r_{x_1 x_2}\sigma_x{}^2 - 2r_{xy}\sigma_x\sigma_y$$

and

$$r_{d_1 d_2} = \frac{r_{y_1 y_2}\sigma_y{}^2 + r_{x_1 x_2}\sigma_x{}^2 - 2r_{xy}\sigma_x\sigma_y}{\sigma_x{}^2 + \sigma_y{}^2 - 2r_{xy}\sigma_x\sigma_y}$$

which is the desired reliability. The necessary constants are the standard deviations of x and y, their respective reliabilities, and the correlation between x and y. In many cases the reliabilities of x and y will be identical and their standard deviations closely alike. Assuming equal reliabilities and standard deviations the formula reduces to

$$r_{d_1 d_2} = \frac{r_{12} - r_{xy}}{1 - r_{xy}}$$

It is clear from this formula that the reliability with which individual differences in increments can be measured is in part a function of the reliability of the original gross scores and in part a function of the extent of the individual differences in the increments or changes which take place during the interval. In general the reliability of measurement of the increment ($r_{d_1 d_2}$) is lower than the reliability of the original gross scores (r_{12}) except in the rare instances where the correlation between the initial and second measurement (r_{xy}) is negative. If the correlation between the initial and second measurement is as high as the reliability of the gross scores then there are no true individual differences in the increments and the reliability with which they can be measured is zero. In general the correlation between the initial and second measurement should be lower as the interval between them increases, hence the reliability with which the increments can be measured should increase with the length of the interval.

BOOK REVIEWS

F. S. FREEMAN. *Individual Differences; the Nature and Causes of Variations in Intelligence and Special Abilities.* New York: Henry Holt & Co., 1934, pp. XI + 355.

After reading this test the impression I received was very similar to that which I got from Jenning's "The Biological Basis of Human Nature." There was not, so far as I could discover, a single statement that was wrong or for which evidence was not adduced, yet the book as a whole gave a decidedly wrong impression. Jennings was justified in writing a text which would be an antidote, as it were, to such books as Wiggam's "Fruit of the Family Tree," but at the end he left the impression that somehow or other, environment was able to add factors or powers or capacities to an organism inherently lacking them. If this were so, then environment is all powerful and stock of little moment. All that we have to do to reach the millenium is to improve environment. There is no reliable evidence for this point of view. Contrary-wise, the evidence seems to indicate the wisdom of stressing stock for long-time improvement; environment for immediate short-time improvement.

Freeman's book is a new departure in the rather badly cultivated field of Individual Differences. He has read widely and assembled his material in excellent shape. Some of the topics treated are—The Extent of Individual Differences; the Influence of Inheritance; the Influence of Environment; the Influence of Race and Nationality; the Differences due to Sex; the Factor of Age; and the Factors of Physical Development and Personality. The trouble with Individual Differences is their number and variety, but Freeman has succeeded admirably in shepherding them into a reasonable number of folds.

As stated previously, Freeman inclines to the environmentalist interpretation of the data at his command. He even relegates the excellent study of foster children by Barbara Burks to a footnote, but devotes five pages of text to the corresponding Chicago study, apparently because the Chicago findings fit his general point of view more closely. Yet in the opinion of many who are competent to judge the California study was the superior one.

This, however, may be regarded as quibbling. What Freeman has done, and done very well indeed, is to assemble a vast amount of apparently unrelated material, analysed it, and made it available for others. Presumably Dr. Freeman is still a young man, but if he

continues as he has started, he should make a very famous name for himself in the field of Educational Psychology. We are sincerely grateful to him.

PETER SANDIFORD.
University of Toronto.

J. G. BEEBE-CENTER. *The Psychology of Pleasantness and Unpleasantness.* New York: D. Van Nostrand Co., 1932, pp. 427 + VIII.

The psychology of pleasantness and unpleasantness is on the face of it a topic which concerns practically every human being. Experimental studies of this subject were started by Fechner in 1871. Many contributions to the study have been made since that time, but the knowledge is scattered in many journals and is not easy to get at. This book in a sense represents the first attempt to gather and organize this knowledge from the many sources into the confines of a single volume. Individuals who are at all familiar with philosophy as well as psychology are acquainted with the fact that the problem of pleasantness and unpleasantness, particularly as it is related to happiness, has concerned philosophers of all ages. Justified, therefore, is the statement made by the author in his introduction to the effect that "an adequate understanding of pleasantness and unpleasantness would constitute not merely an advance of psychology, but a contribution to human welfare." To what extent it would represent a contribution to either of the foregoing would depend upon the author's biases and selection of materials as well as his interpretation of them. To what extent the author succeeds in making contributions to either of the foregoing points of interest will appear in the following lines.

The book begins with a definition of pleasantness and unpleasantness and of the quantitative variable which he calls hedonic tone, or, as it is more frequently referred to in literature, the affective tone. Hedonic tone is to him a quantitative variable having positive and negative values which correspond to pleasantness and unpleasantness. In Chapter II he describes different methods of experimentation used in work on pleasantness and unpleasantness. A large portion of the chapter is given over to a detailed description of Külpe's classification. Chapters III to X inclusive deal with the relationship of hedonic tone to such factors as mental elements, external stimuli, motivating factors, maturation and learning, muscular and glandular responses, memory and nervous processes. The last chapter is devoted to a discussion of his own theory of hedonic tone. Each chapter is followed by very

comprehensive footnotes and references, and the book is indexed by subjects as well as names.

The author has made a very conscious attempt to present the results of studies he reports in their experimental settings and to separate conclusions from results. In this effort he succeeds to a fair extent. However, one familiar with the literature will have no great difficulty in placing Dr. Beebe-Center in terms of methods, attitudes and interpretations that he approves and disapproves. Specifically, for example, the material selected for inclusion in the book as well as theoretical and methodological emphases made by him are much closer to the concerns of so-called structural psychologists than to social, clinical or educational psychologists. By far the most frequent references, for example, are made to the writings of Wundt, Tichener, Nafe, Fechner and P. T. Young. Which is to say that in spite of his ability to observe the importance of the topic to human welfare as well as to psychology, the book as he has written it is of the greatest value to psychologists who tend to work on what may be called the Tichener level of interest. Not much of it will be of interest to the psychologist of more practical concerns who actually is more interested in applying psychology for human welfare. As such, the topic of the relationship of hedonic tone to personality adjustments is not mentioned.

A glimpse of the author's color of life as well as his concept of the role of hedonic tone in life can be gained from the following summary statement of his own theory of the hedonic tone, which he calls "an heuristic hypothesis": "Hedonic tone depends upon a specific type of process in sense-organs, namely, that which, under sensory instructions or their equivalent, mediates bright and dull pressure. When this type of process occurs under hedonic instructions or their equivalent, it gives rise to relative or absolute hedonic tone. Relative hedonic tone depends directly upon the relative density of the process, greater pleasantness being mediated by a change of density which, under sensory instructions or their equivalent, would mediate greater brightness. Absolute hedonic tone likewise depends upon relative density, but in a less direct way. Absolute pleasantness depends upon a density differing from the average for similar past situations in a manner adequate to occasion greater pleasantness; absolute unpleasantness depends upon such a relative density adequate to occasion lesser pleasantness. The process in question is invariably propriocep-

tive. It is aroused by external stimuli—*e.g.*, visual stimuli—indirectly through the muscular adjustments which they bring about."

H. MELTZER.

Psychological Service Center Saint Louis, Missouri.

HARRIET BABCOCK. *Dementia Praecox*. Lancaster: The Science Press, 1933, pp. 167.

The methods and findings reported by Dr. Harriet Babcock in *Dementia Praecox* are significant. They are significant, in the first place, because they do throw additional light on the nature of the abnormality studied; in the second place because the book is an interesting illustration of the possibility of using psychological tests for investigating problems of abnormality in a more specific and intelligent manner than the automatic use generally made by testers not sufficiently informed about the processes or personalities they investigate. With her interpretations of her results, however, few clinical psychologists or psychiatrists will agree.

The thesis of the author is that her findings indicate that dementia praecox is, as Kraeplin originally defined it, a well defined disease entity characterized by gradual deterioration. She is aware of the fact that the psychogenic viewpoint has to a large extent replaced the Kraeplin concept. In her introductory chapter she warns us that psychogenic theories have been abused by persons ignorant of medicine, ignorant of psychology, by those who know medicine and practice psychology and by those who have "little learning" in both fields. It is her opinion that "too often it is assumed that praecox patients have normal ability and that failures to recall are due to repressions, whether conscious or unconscious."

Unlike the general procedure used by mental testers in psychopathic hospitals and mental disease institutions, Dr. Babcock has not used a general battery of tests but instead has developed a method for differentiating an individual's level of abstract intelligence and his present mental efficiency. The discrepancy between these two measures she calls the efficiency index. The tests she used for measuring mental efficiency include such items as the following: Personal information, immediate reproduction of a paragraph, substitution, learning paired-associates, naming days and months in reverse order, drawing designs from memory, immediate reproduction of sentences. In her quantitative comparisons between normal individuals and praecox patients, subjects are divided into sub-groups and equated

for mental level. Subjects compared in this study are two hundred sixteen non-psychotic persons and two hundred six dementia praecox cases. Quantitative results are reported in terms of medians, Q's and PE's. A sixty page appendix gives case studies as well as detailed statistical analyses of the efficiency indices of different types of praecox cases.

Some of the specific conclusions of this study are: There is true memory defect in dementia praecox which differs from normal ability in degree, and which is distinguishable from senile inability by the fact that new data can be fixated, though the process takes longer than it does for normal subjects; most praecox patients were probably mentally impaired at so early an age that it may have interfered with their school achievement; the course of mental deterioration in praecox cases can be quantitatively determined by the method described. Her general conclusion is that, psychologically considered, "dementia praecox is a mental condition in which the past increasingly fails to function effectively in a new present."

For her contribution to the use of psychological tests in a more revealing manner in the study of abnormalities, clinical psychologists owe the author a debt. In the opinion of the reviewer, this debt might have been a much larger one if the author had not seen fit to limit her method and actually handicap its effectiveness by logic-tight thinking concerning the concept of dementia praecox. Specifically, more might have been revealed if similar methods were used not only for studying discrepancies in the intellectual functioning of subjects, but also in making attempts quantitatively to measure emotional factors, even such tabooed things as repressions.

H. MELTZER.

Psychological Service Center, St. Louis.

EDWIN BURKET TWITMYER AND YALE SAMUEL NATHANSON. *Correction of Defective Speech.* Philadelphia: P. Blakiston's Sons & Co., 1932, pp. XVIII + 413.

SARA M. STINCHFIELD. *Speech Disorders.* New York: Harcourt, Brace & Co., 1933, pp. XIII + 341.

Much confusion still remains in the field of speech disorders. There is room for comprehensive books in the field. The books by Dr. Stinchfield and Doctors Twitmyer and Nathanson purport to be comprehensive. The former purports to be a comprehensive psychological study of speech defects; the latter, as its subtitle indicates,

purports to be a "complete manual of psycho-physiological technique for the treatment and correction of defects of speech." As far as purpose is concerned, the two books are in a sense complementary. However, in actuality they are not. Why they are not will appear in the following paragraphs concerning each book.

The book by the University of Pennsylvania authors, though it represents an attempt to offer a technique for the correction of all types of speech defects from faulty articulation of the persistent "baby talker" to cases of profound stammering or stuttering, is relatively simple in its make-up, unified in its organization and consistent in its viewpoint. In fact, more than three hundred of the four hundred thirteen pages of the book are devoted to a description of the corrective material and technique which they advocate in treating speech cases. This portion of the book includes all exercises they consider relevant for treatment, which are: Exercises in body coordination, breathing, vocalization, exercises for peripheral organs, exercises for articulation and integrated speech. The description of the latter represents the bulk of the chapter on materials and techniques.

This treatment is consistent with their general therapy scheme, which views speech production as dependent on three processes: First, correct breathing; second, correct kinesthetic images (Oratans); and third, a correct combination of these two.

The only difficulty in the use of this well constructed manual is that it probably will work for those who have been trained at the University of Pennsylvania under Dr. Twitmyer. Others will find it difficult to use this method in any such exclusive manner. There are other schemes of therapy, and for many cases in the hands of other clinicians it is obvious that other methods can be more effective. However, workers in the field can be thankful to Doctors Twitmyer and Nathanson for their presenting in a very clear-cut and easily intelligible manner one comprehensive scheme. The authors suggest, with some degree of justification, complete adherence to their technique. Only workers who insist upon being consistently logical, however, will have to follow this policy of take-it-all-or-leave-it-all. Adaptations are possible.

Unlike the foregoing book, Dr. Stinchfield's book, particularly Part I, is almost encyclopediac in nature. The two parts into which the book is divided are only loosely related. Part I, called "The Nature of Speech Disorders," is composed of seven chapters dealing

largely with genetic and medical foundations for considering problems of speech disorders. The classification of disorders of speech given in the book is the one recommended by the American Society for the Study of Disorders of Speech. The terms used as well as the general nature of the discussion of the meaning of these terms will make parts of this book very difficult reading for some speech teachers who otherwise could have profited from a reading of the book. The chapter on pathological conditions affecting the speech function includes an almost complete list of all the medical ailments that may in any way be related to speech. Inclusiveness is sometimes gained herein at the expense of clarity and relevancy.

In Part II is reported a statistical study of the speech of three hundred college women and of public school groups. Findings are reported from the use of the Thurstone Personality Schedule, the Jasper-Travis questionnaire on handedness, results of reading tests and speech data sheets. A survey of the book gives one the impression that the author is well informed about her topic and has done some significant work in the field. In its present form the book should serve as a useful book for experts in the field to have as well as a useful aid for orienting students in the field.

H. MELTZER.

Psychological Service Center, St. Louis.